

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:36:19 ; Search time 82 Seconds

(without alignments)
118.183 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378
Sequence: 1 MCRONCRAKSSPEEVISTDE.....INKNVNDKNNSRDYIK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2370	99.7	445	10 Q8W5B2	Q8W5B2 arabidopsis
2	2285	96.1	440	10 Q94CF5	Q94CF5 arabidopsis
3	923	38.8	631	10 Q93V59	Q93V59 arabidopsis
4	615	25.9	223	10 Q9LW50	Q9LW50 arabidopsis
5	552	23.2	107	10 Q8W5B1	Q8W5B1 arabidopsis
6	534.5	22.5	632	10 Q9ZGPO	Q9ZGPO arabidopsis
7	510.5	21.5	692	10 Q9ZNT9	Q9ZNT9 arabidopsis
8	451.5	19.0	186	10 Q94IE6	Q94IE6 arabidopsis
9	423	17.8	851	10 Q23524	Q23524 arabidopsis
10	322.5	13.6	623	10 Q23521	Q23521 arabidopsis
11	266	11.2	295	10 Q23526	Q23526 arabidopsis
12	221	9.3	739	4 Q96BD9	Q96BD9 homo sapien
13	221	9.3	803	4 Q15022	Q15022 homo sapien
14	219	9.2	855	5 Q8T9D8	Q8T9D8 drosophila
15	219	9.2	900	5 Q9NUG9	Q9NUG9 drosophila
16	219	9.2	955	5 Q9W55	Q9W55 drosophila

17	180	7.6	388	10 Q9LU51	Q9LU51 arabidopsis
18	168	7.1	1231	5 Q97140	Q97140 dictyostell
19	156.5	6.6	369	5 Q8T1D9	Q8T1D9 dictyostell
20	153.5	6.5	1309	5 Q8T2H9	Q8T2H9 dictyostell
21	147.5	6.2	4550	5 Q77336	Q77336 plasmodium
22	143	6.0	530	5 Q8R2A0	Q8R2A0 dictyostell
23	141.5	6.0	898	5 Q8T1G7	Q8T1G7 dictyostell
24	140.5	5.9	548	5 Q8T1S2	Q8T1S2 dictyostell
25	140.5	5.9	1204	5 Q8T134	Q8T134 dictyostell
26	140.5	5.9	1711	5 Q77322	Q77322 plasmodium
27	139.5	5.9	1844	5 Q97287	Q97287 plasmodium
28	138	5.8	682	5 Q9GPR8	Q9GPR8 dictyostell
29	136	5.7	855	5 Q8SSV6	Q8SSV6 dictyostell
30	136	5.7	1318	5 Q95PH4	Q95PH4 dictyostell
31	136	5.7	1887	5 Q8SSS8	Q8SSS8 dictyostell
32	135	5.7	164	10 Q9XER7	Q9XER7 oryza sativ
33	135	5.7	1208	5 Q97101	Q97101 dictyostell
34	134	5.6	1245	5 Q8T868	Q8T868 dictyostell
35	134	5.6	1529	5 Q9GQC2	Q9GQC2 dictyostell
36	133.5	5.6	1364	5 Q8T223	Q8T223 dictyostell
37	133.5	5.6	3848	5 Q76737	Q76737 dictyostell
38	133	5.6	1903	5 Q9U0N7	Q9U0N7 plasmodium
39	131.5	5.5	1245	5 Q96195	Q96195 plasmodium
40	131.5	5.5	2472	5 Q8T2M5	Q8T2M5 dictyostell
41	131	5.5	340	5 Q8T212	Q8T212 dictyostell
42	131	5.5	1407	5 Q8T1T5	Q8T1T5 dictyostell
43	131	5.5	2994	5 Q95ZG5	Q95ZG5 dictyostell
44	130.5	5.5	699	5 Q77318	Q77318 plasmodium
45	130.5	5.5	721	5 Q8T218	Q8T218 dictyostell

ALIGNMENTS

RESULT 1	
Q8W5B2	PRELIMINARY; PRT; 445 AA.
AC Q8W5B2:	
DT 01-MAR-2002 (TREMBLrel. 20, Created)	
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE Vernalization 2 protein.	
GN VRN2.	
OS Arabidopsis thaliana (Mouse-ear cress).	
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC eucotids II; Brassicales; Brassicaceae; Arabidopsi	
OX NCBI_TaxID=3702;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=21575875; PubMed=11719192;	
RA GenDall A.R., Levy Y.Y., Willson A., Dean C.;	
RT "The VERNALIZATION2 (VRN2) gene mediates the epigenetic regulation of	
RT vernalization in Arabidopsis.";	
RL Cell 107:525-535(2001).	
DR EMBL: AF284500; AAL32135.1; -	
DR InterPro: IPR000822; Znf_C2H2.	
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.	
DR PROSITE: 51177 MW; 30A3391CE98B877C CRC64;	
SO SEQUENCE	
Query Match	99.7%; Score 2370; DB 10; Length 445;
Best Local Similarity	99.8%; Pred. No. 5.2e-188;
Matches	444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 MCRONCRAKSSPEEVISTDENLLYCKVRLYNIFHLRLSNPFLPCLNYKIGAKRR 60
DB	1 MCRONCRAKSSPEEVISTDENLLYCKVRLYNIFHLRLSNPFLPCLNYKIGAKRR 60
QY	61 KSRSTGMVYFNKCNNTLQKTEVREDCSCPCSLGSGFKGLQPHLNSSHDLEEFERKL 120
DB	61 KSRSTGMVYFNKCNNTLQKTEVREDCSCPCSLGSGFKGLQPHLNSSHDLEEFERKL 120
QY	121 FEETQTVNVSVKLSFIEEGSDDDKEPEPSLCKPKRRKRGGRNNTRLKVCFLPLD 180

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshida N., Yanai Y.;
 RT "Arabidopsis, embryonic flower 2 like 2 gene";
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB053265; BAB58960.1; -
 SO SEQUENCE 186 AA; 21754 MW; D629EB63E9381853 CRC64;

Query Match 19.0%; Score 451.5; DB 10; Length 186;
 Best Local Similarity 50.8%; Pred. No. 1.3e-29;
 Matches 99; Conservative 32; Mismatches 43; Indels 21; Gaps 7;

QY 237 LTSEAVVPAATKTRKLSAESEARSHLL-LQKQFYHSHVOPMALEQVMSDSDSEVD- 294
 DB 6 LTTEAKVP-----AKRSKATSHYLPKHKQFYHSTRGQPLSLQVMSDSDSEVDVK 57
 QY 295 -DDVADFEEDROMLDFVDVKN-DEKQFMHLNMFYRKQVADGHSNACEAFSPRYEKE 352
 DB 58 NDDAAHLESQMLNSMDENIYAEFFIKLNSFVKQRIYADAHIPNACEAFSLHIOE 117
 QY 333 LHRYSLEFWCRLFLIKLNMHGLVDSATINNCNTLLECRNSDPTTNNNSVDPSDS 412
 DB 118 LRSNLSLDCWRFQKMDYGLDRVTMKNKCNFTIYH--NISTNDINDINNNT-----R 170

QY 413 MNNNNNIYDHPNDIN 427
 DB 171 TTDNMDVVD--DDIN 183

RESULT 9
 023524 PRELIMINARY; PRT; 851 AA.
 AC 023524; 023525;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE - Hypothetical 96.8 kDa protein.
 GN ATG16840.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Puidomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chaiwatzis N.;
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: 297342; CAB10457.1; -
 DR EMBL: AL161545; CAB80955.1; -
 DR InterPro: IPR002885; PPR. 9.
 DR Pfam: PF01535; PPR. 9.
 DR TIGRFAMs: TIGR00756; PPR. 10.
 KW Hypothetical protein.
 SO SEQUENCE 851 AA; 96849 MW; 9BF5E46E5FEFEB71C CRC64;

Query Match 17.8%; Score 423; DB 10; Length 851;
 Best Local Similarity 39.6%; Pred. No. 2e-26;
 Matches 101; Conservative 2; Mismatches 0; Indels 152; Gaps 2;

QY 43 PSFLPCLNKTGAKRRK-----SRS 64
 DB 701 PSFLPCLNKTGAKRRKRYAFLLRCSCHSDMLFILLIWKLMNLTSLISRYGHNWMSRS 760

QY 65 TGMVENVKDCNNLTOKTEVRDSCSPFCMSLGGSPKGLQPHLNSHDLFEFEFLFEY 124
 DB 761 TGMVENVKDCNNLTOKTE----- 779
 QY 125 QTVNVSVKLNSPFIPEEGSDDKFPFPPSCSKPRKRRKRGNNRRLKCPPLDSPSL 184
 DB 780 -----GNRGGLYPEATELAGQFEMTSNIPAIASIDAGAKVILTTEAVVP 779
 QY 185 TNGTNGITLLDNGNRGLGYPATLACGFMTSNIPAIANSIDAGAKVILTSEAVVP 244
 DB 780 -----GNRGGLYPEATELAGQFEMTSNIPAIASIDAGAKVILTTEAVVP 826
 QY 245 ATKTRKLSAESEAR 259
 DB 827 ATKTRKLSAESEAR 841

RESULT 10
 023521 PRELIMINARY; PRT; 623 AA.
 AC 023521;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE Hypothetical 70.1 kDa protein.
 GN ATG16810.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Puidomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chaiwatzis N.;
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: 297342; CAB10454.1; -
 DR EMBL: AL161545; CAB80952.1; -
 KW Hypothetical protein.
 SO SEQUENCE 623 AA; 70062 MW; 27174171E2C3F03D CRC64;

Query Match 13.6%; Score 322.5; DB 10; Length 623;
 Best Local Similarity 43.4%; Pred. No. 2.9e-18;
 Matches 76; Conservative 28; Mismatches 36; Indels 33; Gaps 6;

QY 237 LTSEAVVPAATKTRKLSAESEARSHLL-LQKQFYHSHVOPMALEQVMSDSDSEVD- 294
 DB 6 LTTEAKVP-----AKRSKATSHYLPKHKQFYHSTRGQPLSLQVMSDSDSEVDVK 57
 QY 295 -DDVADFEEDROMLDFVDVKN-DEKQFMHLNMFYRKQVADGHSNACEAFSPRYEKE 352
 DB 58 NDDAAHLESQMLNSMDENIYAEFFIKLNSFVKQRIYADAHIPNACEAFSLHIOE 117
 QY 333 LHRYSLEFWCRLFLIKLNMHGLVDSATINNCNTLLECRNSDPTTNNNSVD 407
 DB 118 LRSNLSLDCWRFQKMDYGLDRVTMKNKCNFTIYH--NISTNDINDINNNT-----R 151

RESULT 11
 023526 PRELIMINARY; PRT; 295 AA.
 AC 023526;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)

Hypothetical 33.6 kDa protein.
 GN ATG16850.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Beyer M., Stiekema W., Murphy G., Wandutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Pudumench P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chalevalais N.,
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z97342; CAB10459.1;
 DR EMBL; AL161545; CAB80956.1;
 KW Hypothetical protein.
 SQ SEQUENCE 295 AA; 33564 MW; F22BD5E82031423 CRC64;

Query Match 11.2%; Score 266; DB 10; Length 295;
 Best Local Similarity 53.4%; Pred. No. 5, 3e-14;
 Matches 70; Conservative 3; Mismatches 30; Indels 28; Gaps 4;

278 MALEOVMSDSEDDVDVADVEDRQ-----MDDFVYVNDKDEQ 318
 DB 1 MALEOVMSDSEDDVDVADVEDRQVHDFPLNSLQOQKMYVMDLDDVDVNDKDEQ 60
 QY 319 FMHLMNSFVRKQVADGHSWACEAFSFRYEKEHRRYSSEFWCRFLIKLMHGLVDS 378
 DB 61 FMHLMNSFVRKQVADGHSWACEAFSFRYEKEHRRYSSEFWCRFLIKLMHGLVDS 111
 QY 379 ATINNCNTILE 389
 DB 112 ALANVHTRE 122

RESULT 12
 Q96BD9 PRELIMINARY; PRT; 739 AA.
 AC Q96BD9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Joined to JAZF1
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Strauberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015704; AA015704.1;
 DR InterPro; IPR000822; Znf_C2H2.
 DR SMART; SM00355; Znf_C2H2.1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW DNA-binding; Zinc-finger.
 SQ SEQUENCE 739 AA; 83054 MW; A8B30EBC3FD38D56 CRC64;

Query Match 9.3%; Score 221; DB 4; Length 739;
 Best Local Similarity 21.1%; Pred. No. 9e-10;
 Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

58 RRRK-----SRSTGVVFNKDCNNTLOKTEVEREDSCSPGSMGSGFKGLQFHLNNSHD 112
 DB 413 RRRKDPENRQKRLRYOFLYNNNTROOTEARDDLHCPMCTLNCRKLYSLKHLKLCHS 472

113 LFEFEFKLFEEQTVNVSVKLNSTFEEESDDDKFEPFSLCSNPK-RROR---GRNN 168
 DB 473 RFLFNVYHPKARLDVSI-----NECIDG-----SYAGNPQDIHQPFANSRNG 518
 QY 169 -TRRLKVCPLPDSLSLNGTENGITLLDNGRGLGPEATELAGOFEWTSNIPALIAS 227
 DB 519 PVKRFPIHTLVCRRKFKKASSEFLSESD----- 549
 QY 228 SLDAKAVILITSEAVVPATKTRKLSAERSEARSHLLQKROFYSHRVOPMALQVMSDR 287
 DB 550 -----EVEQQTYSNG-----NLYTHSDTCLPLRPOEM--EV 581
 QY 288 DSEDEVDVADVEDRQMDLDFVYVNDKQFMHLMNSFYKQVADGHSWACEAFSR 347
 DB 582 DSEDEKPEMLREKTIQIEEFSVNEGEKRVKMLNLYHKRGITADNOMHACMLFEVE 641
 QY 348 FYEKEHRSYSLFCWCRFLIKLMHGLVDSATINNCNTILENCRRSSPTTTNNNSVD 407
 DB 642 NYGOKRIK-KNLGRFMHLVSMHDFNLISMSIDKAVTKLRMQ-----QKLE 689
 QY 408 RPSDSNTNNNTVDDHPN-----DINKNNVNDKONNS 439
 DB 690 KGSASPAHEETTEBQNGTANGFSEINSKEALETDSVS 728

RESULT 13
 Q15022 PRELIMINARY; PRT; 803 AA.
 ID Q15022;
 AC Q15022;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE KIAA0160 protein (Fragment).
 GN KIAA0160.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Ishiava K., Nomura N.;
 RT Prediction of the coding sequences of unidentified human genes. IV.
 RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
 RT analysis of cDNA clones from human cell line KG-1.";
 RL DNA Res. 2:167-174(1995).
 DR EMBL; D63861; BA009931.1;
 DR InterPro; IPR000822; Znf_C2H2.
 DR SMART; SM00355; Znf_C2H2.1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 803 AA; 89963 MW; CDFB901A35F29A7C CRC64;

Query Match 9.3%; Score 221; DB 4; Length 803;
 Best Local Similarity 21.1%; Pred. No. 1e-09;
 Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

58 RRRK-----SRSTGVVFNKDCNNTLOKTEVEREDSCSPGSMGSGFKGLQFHLNNSHD 112
 DB 477 RRRKDPENRQKRLRYOFLYNNNTROOTEARDDLHCPMCTLNCRKLYSLKHLKLCHS 536
 QY 113 LFEFEFKLFEEQTVNVSVKLNSTFEEESDDDKFEPFSLCSNPK-RROR---GRNN 168
 DB 537 RFLFNVYHPKARLDVSI-----NECIDG-----SYAGNPQDIHQPFANSRNG 582
 QY 169 -TRRLKVCPLPDSLSLNGTENGITLLDNGRGLGPEATELAGOFEWTSNIPALIAS 227
 DB 583 PVKRFPIHTLVCRRKFKKASSEFLSESD----- 613
 QY 228 SLDAKAVILITSEAVVPATKTRKLSAERSEARSHLLQKROFYSHRVOPMALQVMSDR 287
 DB 614 -----EVEQQTYSNG-----NLYTHSDTCLPLRPOEM--EV 645

QY 288 DSEDEVDVADFEEDROMLDFVDVYKDEKQFMHLSNFSVRKORVIADGHISWACEAFGR 347
 DB 646 DSEDEKDEPMLEKERTITQIEESDVNEGKEVMKLMNLHVMHGFADQNHACHLFEVE 705
 QY 348 EYEKELHRYSLFPCWRLFLIKMNHGLVDSATINNCNTILENCRNSDPTTTNNNSVD 407
 DB 706 NGOKRIK-KNLGRNFMHLMVSHMDFNLISIMSKAVYKLEMO-----QKLE 753
 QY 408 RPSDSTNNNNIVDHPN-----DINNKNVNDKDNNS 439
 DB 754 KGESASPAHEITTEBONGTANGFSEINSKEALETDSVS 792

RESULT 14

Q8T9D8 PRELIMINARY; PRT; 855 AA.
 AC Q8T9D8; 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE SD04959p.
 GN SU(2)12.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Farfan D., Frisze E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rudin G.M., Gelinker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY069809; AAL39954.1;
 SQ SEQUENCE 855 AA; 95316 MW; C10FEC6013954105 CRC64;

Query Match 9.2%; Score 219; DB 5; Length 855;
 Best Local Similarity 22.6%; Pred. No. 1.6e-09;
 Matches 74; Conservative 47; Mismatches 113; Indels 94; Gaps 8;

QY 68 VFNFKDCNNTLOKTEVEDSCSPFCMCGSFGKLGPHLNSHDLFEFEFKLFEEYQTV 127
 DB 391 IYVNMYSNNTROQETIYQELNCPWCGLDCLRLYLKHLKCHARFNTTYPAGSGARI 450
 QY 128 NVSVKLSNFIIEEGSDDDKEP-----FSLCSKPKRRKRGGRN--NTRRLKVC--FLP 178
 DB 451 DVTIN-DAVDGSYAGSPYDLAPGSSSFARTCGPVRRSVTSIMVCRPRKTCIDEFLE 509
 QY 179 LDSPLTNGTENGITILLDNGNGLGCPATELAGQFEMTSNIPPAIHSLSLDAKAVIIT 238
 DB 510 LDEDEISN----- 517
 QY 239 SEAVVPATRTKRLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSDEYDDVA 298
 DB 518 -----GRSYITGHNRLLYHHTETCLPVHPKEL--DIDSGESDPLWL 556
 QY 299 DEEDROMLDDFVDVYKDEKQFMHLSNFSVRKORVIADGHISWACEAF-----SREYEKEL 353
 DB 557 ROKTIOMIDEFSDVNEGEKELKLMNLHVMRHGFVGDCLPIACEMFLDAKGTETVIRKNL 616
 QY 354 HRYSSLFPCWRLFLIKMNHGLVDSATI 381
 DB 617 YNFTLHMC-----SLFDYGLIAAETV 638

RESULT 15

Q9NUG9 PRELIMINARY; PRT; 900 AA.
 AC Q9NUG9; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE SU(2)12. OR CG8013.
 GN SU(2)12 OR CG8013.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Birve A., Rasmuson-Lestander A., Larsson J.,
 RT "Suppressor of zeste 12.";
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF149047; AAF73149.1;
 DR FlyBase; FBgn020867; Su(2)12.
 DR InterPro; IPR000822; Znf_C2H2.
 DR SMART; SM00355; Znf_C2H2; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 SQ SEQUENCE 900 AA; 100103 MW; 53BA0D83C49EC92F CRC64;

Query Match 9.2%; Score 219; DB 5; Length 900;
 Best Local Similarity 22.6%; Pred. No. 1.7e-09;
 Matches 74; Conservative 47; Mismatches 113; Indels 94; Gaps 8;

QY 68 VFNFKDCNNTLOKTEVEDSCSPFCMCGSFGKLGPHLNSHDLFEFEFKLFEEYQTV 127
 DB 391 IYVNMYSNNTROQETIYQELNCPWCGLDCLRLYLKHLKCHARFNTTYPAGSGARI 450
 QY 128 NVSVKLSNFIIEEGSDDDKEP-----FSLCSKPKRRKRGGRN--NTRRLKVC--FLP 178
 DB 451 DVTIN-DAVDGSYAGSPYDLAPGSSSFARTCGPVRRSVTSIMVCRPRKTCIDEFLE 509
 QY 179 LDSPLTNGTENGITILLDNGNGLGCPATELAGQFEMTSNIPPAIHSLSLDAKAVIIT 238
 DB 510 LDEDEISN----- 517
 QY 239 SEAVVPATRTKRLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSDEYDDVA 298
 DB 518 -----GRSYITGHNRLLYHHTETCLPVHPKEL--DIDSGESDPLWL 556
 QY 299 DEEDROMLDDFVDVYKDEKQFMHLSNFSVRKORVIADGHISWACEAF-----SREYEKEL 353
 DB 557 ROKTIOMIDEFSDVNEGEKELKLMNLHVMRHGFVGDCLPIACEMFLDAKGTETVIRKNL 616
 QY 354 HRYSSLFPCWRLFLIKMNHGLVDSATI 381
 DB 617 YNFTLHMC-----SLFDYGLIAAETV 638

Search completed: June 13, 2003, 15:45:22
 Job time : 85 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:17:53 ; Search time 1928 Seconds

(Without alignments)
3738.066 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRACKSPEVISTDE.....INNNNNNDKNNNSRDYIK 445

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+P2n_model -DEV=x1h
-O=/cgn2.1/USPTO_spool/US09890220/runat_13062003_144851_16490/app-query.fasta.1.583
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -IOFCL=0 -LOOEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09890220_ECGN_1.1.1906_Erunat_13062003_144851_16490 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	778	32.7	672	10	AV822548
2	760.5	32.0	1128	11	AV104964
3	568	23.9	680	10	AM038171
4	553	23.3	769	14	BO505017
5	545	22.9	846	12	BG648271
6	528	22.2	480	9	AJ469021
7	510.5	21.5	561	14	BQ740672
8	509	21.4	600	13	B1479743
9	506	21.3	391	10	AV817525
10	498.5	21.0	569	14	BO611550
11	456.5	19.2	474	10	AM234600
12	450	18.9	467	12	BG580495
13	448	18.8	405	12	BG314044
14	444.5	18.7	662	14	BQ279630
15	442.5	18.6	612	10	BE203328
16	437	18.4	444	12	BG551103
17	436	18.3	554	13	B1784785
18	422	17.7	429	14	BQ459233
19	420	17.7	546	13	B1321750
20	396.5	16.7	576	10	AM076157
21	394	16.6	364	9	A1164598
22	386	16.2	732	14	BQ841916
23	375	15.8	561	14	BQ163202
24	374	15.7	644	13	BQ453006
25	372	15.6	757	13	BQ460553
26	367.5	15.5	445	10	AM202144
27	362.5	15.2	594	10	AM000259
28	347	14.6	519	14	BQ168835
29	343.5	14.4	326	9	A1163743
30	329	13.8	727	13	B1932726
31	306	12.9	727	14	BQ801572
32	304.5	12.8	560	9	A1795561
33	304	12.8	526	17	CNS00MYB
34	303	12.7	586	14	BQ243123
35	302.5	12.7	529	12	BG459005
36	299.5	12.6	515	12	BF481866
37	274	11.5	677	14	BQ659737
38	271	11.4	322	10	AM309408
39	269	11.3	236	14	BQ904049
40	268	11.3	804	17	BH587461
41	267	11.2	553	12	BG909756
42	267	11.2	553	14	BQ620168
43	257	10.8	523	13	B1992698
44	256.5	10.8	632	12	BE821501
45	254.5	10.7	416	10	AM432863

ALIGNMENTS

RESULT 1
AV822548
LOCUS
DEFINITION AV822548 RAPL5 Arabidopsis thaliana CDNA clone RAPL05-08-K19 5',
ACCESSION AV822548
VERSION AV822548.1 GI:19864597
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana
thale cress.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 672)
Seki, M., Narusaka, M., Ishida, J., Kamuya, A., Satou, M., Nakajima, M.,

Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arkawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 unpublished (2002)
 Contact: Motoki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msekierc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). This clone is in a
 modified pBluescript vector as a SstI/XhoI insert. Please visit our
 web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
 further details.

FEATURES

source

1. 672
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAF105-08-K19"
 /clone_id="RAF15"
 /dev_stage="rossette plants"
 /lab_host="SOLR"
 /note="Site_1: SstI, Site_2: XhoI; subjected to
 dehydration-treated(1,2,5,10,24 hr)"
 BASE COUNT 197 a 116 c 143 g 216 t
 ORIGIN

Alignment Scores:

Pred. NO.: 6.11e-80 Length: 672
 Score: 778.00 Matches: 144
 Percent Similarity: 99.32% Conservative: 1
 Best Local Similarity: 98.63% Mismatches: 1
 Query Match: 32.72% Indels: 0
 DB: 10 Gaps: 0

US-09-890-220-2 (1-445) x AV822548 (1-672)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 |||||
 DB 233 ATGTGTAGGAGCAATGTGCGCGAATCTCTACCGGAGAGTAGTATTTCACTCATGAG 292
 |||||
 QY 21 AspLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 |||||
 DB 293 AATCTCTGATATATGTAACTGTTCGACATTAACATCTTCACCTTGCGCTCTA 352
 |||||
 QY 41 GlysProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 |||||
 DB 353 GGCACCCATCGTTTCTGCCAGATCTTGAACATAAGATTGGGCAAGGCAAAAGA 412
 |||||
 QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 |||||
 DB 413 AAGTCAGATCTACTCGGAGTGTGATTTCACTTAAGATTGAATTAATCAATTACAA 472
 |||||
 QY 81 LysThrGluValArgLysAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
 |||||
 DB 473 AGAAGTGAAGTGGAGAGATGTTCTGTTCATTTGCTATGCTATGCTAGCTTC 532
 |||||
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
 |||||
 DB 533 AAGGGCTGCAATTTGATTTGATTCATCTCATGATTTAATTGAATTTGAGTTCAAGCTT 592
 |||||
 QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 |||||
 DB 593 TTGGAGAAATACACAGACATTAATGTTCTGTAAACCTTAATTCCTCATATTGAGGAA 652
 |||||
 QY 141 GluGlySerAspAsp 146
 |||||
 DB 653 GAGAGAGTGAATGAT 670

RESULT 2

AY104964
 LOCUS AY104964 1128 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PCO105783 mRNA sequence.
 ACCESSION AY104964
 VERSION AY104964.1 GI:21208042
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays.
 Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 1128)
 Hainey, C.F., Dolan, M., Mao, G.H., Vogel, J.M., Whitsett, M.S.,
 Arthur, L.W., Hanafey, M., Morgan, M., and Tingey, S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1128)
 REFERENCES
 Direct Submission
 Coe, E.C.
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 Location/Qualifiers

FEATURES

source

1. 1128
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /db_xref="maizedb:636128"
 /clone="PCO105783"
 /clone_id="Maize Mapping Project/Dupont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"
 BASE COUNT 355 a 199 c 242 g 332 t
 ORIGIN

Alignment Scores:

Pred. NO.: 1.48e-77 Length: 1128
 Score: 760.50 Matches: 157
 Percent Similarity: 64.35% Conservative: 47
 Best Local Similarity: 49.53% Mismatches: 80
 Query Match: 31.98% Indels: 33
 DB: 11 Gaps: 7

US-09-890-220-2 (1-445) x AY104964 (1-1128)

QY 86 GluAspCysSerCysProPheCysSerMetLeuCysGlySerPheGlyLeuGlnPhe 105
 |||||
 DB 12 GAAGATTTCCTGTCATTTGCTATGACATGTGGAAGCTTCAGGCTGTGAGATGC 71
 |||||
 QY 106 HisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGluTyrGln 125
 |||||
 DB 72 CATTTAACTCATCATGATCATCTCCATTCATGAGTTTGGATATCTGAAGATACAG 131
 |||||
 QY 126 ThrValAsnValSerValLysLeuAsnSerPhe-----IlePheGluGluGluGly 142
 |||||
 DB 132 GTTGTAAATGTTAGTTGAAAGCTCATGCTTGAAGACAGACCTTTTGGCGAG- 185
 |||||
 QY 143 SerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArgGln 162
 |||||
 DB 186 GCGCTGATCCAGAGCATCAACATTTTATTCCTCAAGGTTTAAAGCGTAGACGA 245
 |||||
 QY 163 ---ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSer 181
 |||||
 DB 246 TCAAGAACCAATGAGAAATCAAGCATGACCTCAGATATTATGCAATTCAGTTCA 305
 |||||
 QY 182 ProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGly 201
 |||||
 DB 306 CCTGAAGAGAGGAGGAGATCTGAGACCAACTTTGT-----CAAGGGAGAAATGGG 356

QY 202 LeuGIyTyPrGluAlaThrGluLeuAlaGlyInpHeGluMetThrSerAsnIlePro 221
 Db 356 ----- 356
 QY 222 ProAlaIleAlaHisSerSerLeuAspAlaGlyAlaVal-----IleLeuThr 238
 Db 357 ACTTCTGTGAGCAAAATGCTTCATTCATCTCTGCTCAATCTTACATGAGCAGCAATCTTCA 416
 QY 239 SerGluAlaValAlaProAlaThrIleThrIleGlyLeuSerAlaGluArgSerGluAla 258
 Db 417 CCACCAACAGTACTACAGTTGGGAGAACAGGAGGCTATCT--GAGAGATCTGACCT 473
 QY 259 ArgSerHisLeuLeuGluInLysArgInPheTyHisSerHisArgValGlnPromet 278
 Db 474 AGAATTCGGCAACTCTCCAAAAAGACAGTTCTTCATCTCCACAGGGCCGACCAATG 533
 QY 279 AlaLeuGluInValMetSerAspArgSerGluAspGluValAlaAspAspValAla 298
 Db 534 CAACGTGAGCAAGTGTCTCGGACCGGTAGTGAATGAAATGATGATGATATGCT 593
 QY 299 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGln 318
 Db 594 GACTTCGAGATAGAGATGCTGTGATATTTGTTGATGATGACGAAAGATGAAACTT 653
 QY 319 PheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIle 338
 Db 654 ATTATGATATGTGATTCATTTCTTGTGAAACAAAGCTTGTGATGATGATGATGAT 713
 QY 339 SerTrpAlaGlyGluAlaPheSerArgPheTyGlyLysGluLeuHisArgTySerSer 358
 Db 714 CTTGGGGCTCGAGAGCTTCTCCAGATGATGACGACCAACTATGATCAAAATCTCTCT 773
 QY 359 LeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSer 378
 Db 774 CTGCTGTGGGGTGGCTTCTTCATGATTAACCTTGTGACCAACATTAACATTTAATGCTC 833
 QY 379 AlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
 Db 834 CGCACTATGACACATGCAATACAGTCTCTCAATTTTACAAAGAAAGAAC 884
 RESULT 3
 AM038171 680 bp mRNA linear EST 18-MAY-2001
 LOCUS EST279828 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION clone cLET1N5, mRNA sequence.
 ACCESSION AM038171
 VERSION AM038171.1 GI:5896925
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 REFERENCE 1 (bases 1 to 680)
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Jiang, F., Upton, J.,
 Romling, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nieman, W.,
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni,
 J.,
 Generation of ESTs from tomato leaf tissue
 Unpublished (1999)
 CONTACT CUGI
 JOURNAL Clemson University Genomics Institute
 COMMENT Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 5 prime sequence.
 FEATURES
 Location/Qualifiers
 1..680
 /organism="Lycopersicon esculentum"
 /cultivar="Rio Grande Ptor"
 /db_xref="taxon:4081"
 /clone="cLET1N5"

/clone.lib="tomato mixed elicitor, BTI"
 /tissue.type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XL1-Blue MRF"
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenchion, Etx,
 Okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."
 BASE COUNT 192 a 120 c 164 g 204 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,95e-55 Length: 680
 Score: 568.00 Matches: 107
 Percent Similarity: 84.35% Conservative: 17
 Best Local Similarity: 72.79% Mismatches: 21
 Query Match: 23.89% Indels: 2
 DB: 10 Gaps: 2
 US-09-890-220-2 (1-445) x AM038171 (1-680)
 QY 254 GluArgSerGluAlaArgSerHisLeuLeuGluInLysArgGlnPheTyHisSerHis 273
 Db 4 GAGCGCTGTGATCCCGAAGATGCTGACCTCTCGCAAAAAGCAATCTTCTCATCTCAT 63
 QY 274 ArgValGlnPrometAlaLeuGluInValMetSerAspArgAspSerGluAspGluVal 293
 Db 64 AGGGCCGAGCCCATGCGACCTGAGCAGAGTTTATGAGCAGCAGAGAGTGAATGAT 123
 QY 294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
 Db 124 CATGATGATGTTGCGATCTTGAAGATGAGAGATGCTTGATGATTTGTGATGTGACC 183
 QY 314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
 Db 184 AAAGATGAAAGCAAGTGTGATGCTGTGAGCACTCATTTGTAGAAAGCAAGGTTG 243
 QY 334 AlaAspGlyHisIleSerTrpAlaGlyGluAlaPheSerArgPheTy---GluLysGln 352
 Db 244 GCAGATGTGATATATCCCTTGGCAGTGTGAGGCTTTCAAACTCAGTGCAGGCTT 303
 QY 353 LeuHisArgTySerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372
 Db 304 GCCCAAGCACACGACCATATGCAAGGCTTGTGACATTTATCATATGACGTGTGCAAC 363
 QY 373 HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392
 Db 364 CATGCGCTGTGATGCGCGGTACATTAACATTTAATTAATTAATTAATTAATTAATTA 423
 QY 393 AsnSer---SerAspThrThr 398
 Db 424 AACCAATACAGTATCTACT 444
 RESULT 4
 B0505017 769 bp mRNA linear EST 22-JUL-2002
 LOCUS EST612432 Generation of a set of potato cDNA clones for microarray
 DEFINITION analyses mixed potato tissues Solanum tuberosum cDNA clone STMGB40
 3' end, mRNA sequence.
 ACCESSION B0505017
 VERSION B0505017.2 GI:21920982
 KEYWORDS EST.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 769)
 Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
 Karamycheva, S.A.

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: curesgen.com web site: www.resgen.com

Seq primer: -40RP from G1bco
High quality sequence stop: 426.

FEATURES

source

1..561

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl076-3274"

/clone_1b="Gm-cl076"

/tissue_type="wounded cotyledons"

/dev_stage="11 day old seedlings"

/lab_host="DH10B"

/note="Vector: phuescript II SK+; Site.1: EcoRI; Site.2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene phuescript II SK+ library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the phuescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Lila Vodka Lab University of Illinois)."

BASE COUNT

166 a 91 c 123 g 181 t

ORIGIN

Alignment Scores:

Pred. No.: 7.41e-49 Length: 561
Score: 510.50 Matches: 93
Percent Similarity: 77.78% Conservative: 19
Best Local Similarity: 64.58% Mismatches: 27
Query Match: 21.47% Indels: 5
DB: 14 Gaps: 1

US-09-890-220-2 (1-445) x B0740672 (1-561).

OY 280 LeungluInvalmetSerAspArgSerGluAspGluValAspAspValAlaAsp 299
DB 2 CTGAGAACAGTGTATACAGCCGTGATGAGAACAGAGTGTATGACGACATTCGACAGT 61
OY 300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnAspGlnuysGlnPhe 319
DB 62 CTTGAGAGTAAAGAGATGCTTGACGATTTGTGATCTTCCAAAGATGAAAACAGCTC 121
OY 320 MethisLeuTrpAsnSerPheValArgGlnArgValAlaLeuAspGlnuysIleSer 339
DB 122 ATGCATCTCTGACCTTTTATAGGAGCAAGGCTGTAGACAGTGTATGTCG 181
OY 340 TTPAlaGysGluAlaPheSerArgPheTyrGluysGluLeuHisArgTyrSerSerLeu 359
DB 182 TGGGCGCTGAGGCTTTTCCACCTTCATGAAAGAGCTGATCATCTCCACGCTTA 241
OY 360 PheTyrCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAla 379
DB 242 TTTTGGTGTGGAGTATTCATGATCAAACTTGGATCATGATGCTCTTCTTGATGCTCTG 301
OY 380 ThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThr 399
DB 302 ACAATGACCACTGAGCATAGTATGATGATTCAGAAATACAGGATCGGCTACAAGA 361
OY 400 ThrAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAla 419
DB 362 AAAAAT-----TGAAAGCAATAGGAAACATGACTCAACAAACAGATATA 406

OY 420 ValAspHisPro 423
DB 407 CAGTCTTCTCCG 418

RESULT 8

BI479743

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..600

/organism="Triticum aestivum"

/db_xref="taxon:4565"

/clone="WHE3451_E10_119"

/tissue_type="Wheat pre-anthesis spike cDNA library"

/dev_stage="Adult plant"

/lab_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site.1: EcoRI; Site.2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagemids in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT

165 a 126 c 143 g 166 t

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-48 Length: 600
Score: 509.00 Matches: 102
Percent Similarity: 62.18% Conservative: 18
Best Local Similarity: 52.85% Mismatches: 44
Query Match: 21.40% Indels: 2
DB: 13 Gaps: 2

US-09-890-220-2 (1-445) x BI479743 (1-600)

OY 237 LeuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
DB 21 CTTACACACAGCAGTACAGTGTGGGAAAGACAGAAACTATCTCGGAGGAGT 80
OY 257 GluAlaArgSer----- 260

```

Db      81 GATCCAGAAA-GTACCCCTTCTGTTACATACTCGTATATCATCTAAGCATTTGTA 139
      261 -----HisLeu----- 262
Db      140 TATATAGCTGTTTGGTTTTCAGACATTAAGAAAAATTTAAACCACTGTTTTT 199
      263 -----LeuLeuGlnLysArgGlnPheTyrHisSerHis 273
Db      200 ATGTATATATATATGTCGACCGCCGACAACTTTCGAGAAACGTCAGTCTTCATTTCTCAC 259
      274 ATGValGlnPheMetAlaLeuGlnLysValMetSerAspArgAspSerGlnAspGlnVal 293
      260 AGGCGACAGCAACATGAGTGGACAGATTTGTCAGATCATGACAGCAGATGAGTT 319
      294 AspAspAspValAlaAspPheGlnAspArgGlnMetLeuAspAspPheValAspValAsn 313
      320 GACGATGACATGTCGCGACTTGGAGATGACGATGCTTGCATTTTCTTGATGTCACG 379
      314 LysAspGlnLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
      380 AAAGATGAGAACCTTATTTATGCAATATGGAATTCATTTATTCGGAACAAAGGCTGA 439
      334 AlaAspGlnHisIleSerTrpAlaCysGlnAlaPheSerArgPheTyrGlnLysGlnLeu 353
      440 GCTGATGGGCATATACCTTGGCGCTGTGAGGCTTCTCCGCGCTCATGACCGACGCTT 499
      354 HisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHis 373
      500 GTACAAACCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 559
      374 GlnLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
      560 AGCTTGCTGATGCGCGCGCCGACATGACATGCAACACC 598

```

RESULT 9
LOCUS AV817525/c 391 bp mRNA linear EST 01-APR-2002
DEFINITION AV817525 RAF19 Arabidopsis thaliana CDNA clone RAF109-94-M06 3',
mRNA sequence.
ACCESSION AV817525
VERSION AV817525.1 GI:19859415
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 391)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itch, Y., Ishi, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

FEATURES
SOURCE location/Qualifiers
1..391
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

```

/clone="RAF109-94-M06"
/clone.lib="RAF19"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/node="Site_1: BamHI; Site_2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"
BASE COUNT 102 a 100 c 73 g 116 t
ORIGIN
Alignment Scores:
Pred. No.: 1,43e-48 Length: 391
Score: 506.00 Matches: 89
Percent Similarity: 87.07% Conservative: 12
Best Local Similarity: 76.72% Mismatches: 15
Query Match: 21.28% Indels: 0
DB: Gaps: 0
US-09-890-220-2 (1-445) x AV817525 (1-391)

```

```

      274 ATGValGlnPheMetAlaLeuGlnLysValMetSerAspArgAspSerGlnAspGlnVal 293
      389 CGAGCTCACGCCATGCTGCTTACAGACAGTCTTGGACCGGATAGTGAAGATGAGTT 330
      294 AspAspAspValAlaAspPheGlnAspArgGlnMetLeuAspAspPheValAspValAsn 313
      329 GATGATGATGATGCGAGATTTTGAAGATGAAGATGCTGATGATTTCTGATATGACT 270
      314 LysAspGlnLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
      269 AAAGATGAGAACCTTATTTATGCAATATGGAATTCATTTATTCGGAACAAAGGCTGA 210
      209 GCAGATGTCACATTCATTCATGCGCATGCGCATGCTCAAGATTCGACGACCGACCATCATG 150
      354 HisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHis 373
      149 GTTCGAACCGCCACTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 90
      374 GlnLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleGlnLys 389
      89 GGTCTTCTGATGCCCGAACCATGAAACATGTAATACCTTCTCGAA 42

```

RESULT 10
LOCUS BO611550 569 bp mRNA linear EST 26-JUN-2002
DEFINITION sap61f09.y1 Gm-c1087 Glycine max CDNA clone SOYBEAN CLONE ID:
Gm-c1087-5250.5', similar to TR:092NT9.092NT9
FERTILIZATION-INDEPENDENT SEED 2 PROTEIN.1', mRNA sequence.
ACCESSION BO611550
VERSION BO611550.1 GI:21601219
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 569)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna,
A., Bolla, B., Merritt, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,
Y., Pearson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr,
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,
R., Waterson, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

FEATURES
SOURCE location/Qualifiers
1..569
/organism="Glycine max"
/db_xref="taxon:3702"

(pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 133 a 86 c 107 g 148 t
ORIGIN

Alignment Scores:

Pred. No.:	1.14e-42	Length:	474
Score:	456.50	Matches:	89
Percent Similarity:	78.36%	Conservative:	16
Best Local Similarity:	66.42%	Mismatches:	28
Query Match:	19.20%	Indels:	1
DB:	10	Gaps:	1

US-09-890-220-2 (1-445) x AM234600 (1-474)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
|||||
DB 43 ATGTCCCGCAAAATTCCTCCGTACACCATGCCGGTGAAGAAGAAATTCACCTGATGAG 102
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
:::|||||
DB 103 AGCTTTAAATTTATTCAGACGCTGTGAACGTGTACAAATATCTCTACCGCGCTCTT 162
QY 41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
|||||
DB 163 CAAATTCCTCTTCTTCTTGAAGATGTTGCTTAAATAAGAGCAAGGCGTAAAGG 222
QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnThrLeuGln 80
:::|||||
DB 223 AGGTTGAGAGCA---GGAATGTGTGATTTTCATTTAGGATCAGTACACATTTCTCGG 279
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
|||||
DB 280 AAAAGTGAAGTACGACGAGACCTTTCTGTCCGTTTCTGTGATGATGATGAGAGCTTT 339
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
|||||
DB 340 AAGGTTTGCGATTCATCTTGTTCATCAGATGATGATTCATCTTGAAGTTGAGTTG 399
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
|||||
DB 400 ACTGAAGATTACCAAGCAGTGAATCTCTCCGGAATTAAT 441

RESULT 12
BG580495

LOCUS EST482220 GVN Medicago truncatula cDNA clone pGVN-57B2 5' end, mRNA
DEFINITION

ACCESSION BG580495
VERSION BG580495.1 GI:13595559
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

REFERENCE 1 (bases 1 to 467)

AUTHORS Fedorova,M., Plešon,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Town

C.D., Van Aken,S., Utecher,T., Cho,J. and Fraser,C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago

TITRUNCATULA, 2001

JOURNAL Unpublished (2001)

COMMENT Department of Agronomy and Plant Genetics

University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA

Tel: 612 625 5715
Fax: 651 649-5058

Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M381546e TIGR sequence name: MTCOT01TK More information is available at: <http://www.medicago.org>
Seq primer: SKmod (CTA GAA CTA gta gat CC).
Location/Qualifiers

FEATURES

source

1..467

/organism="Medicago truncatula"

/cultivar="genotype A17"

/db_xref="taxon:3880"

/clone="pGVN-57B2"

/clone_11b="GVN"

/tissue_type="N2-fixing root nodules"

/dev_stage="effective root nodules harvested one month post inoculation with Sinorhizobium meliloti"

/lab_host="E. coli strain X10LR"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from effective root nodules harvested one month post inoculation with Sinorhizobium meliloti. The cDNA was directionally ligated into the Uni-ZAP XR vector from StrataGene and packaged using GigaPack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in X10LR cells."

BASE COUNT 136 a 87 c 95 g 149 t

ORIGIN

Alignment Scores:

Pred. No.:	6.39e-42	Length:	467
Score:	450.00	Matches:	91
Percent Similarity:	70.32%	Conservative:	18
Best Local Similarity:	58.71%	Mismatches:	40
Query Match:	18.92%	Indels:	6
DB:	12	Gaps:	3

US-09-890-220-2 (1-445) x BG580495 (1-467)

QY 17 SerThrAspGluAsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHis 36
:::|||||
DB 3 ACAGCTGATGAAGACCTTTTGATTTATTCGAAACCTGTTGATTTACATATTTCTGAT 62
QY 37 LeuArgSerLeuGlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAla 56
|||||
DB 63 CGTCGTTCTTCAACAATCCATCTTCTTGAAGATGTTGCCGCTATATAAAGCA 122
QY 57 LysArgLysArgLysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsn 76
|||||
DB 123 AAGCAAAAAGAGAGGTCGAGCA---GGAATGTGTGTTTCAACTATAGGACGTTAC 179
QY 77 AsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPheCysSerMetLeu 96
|||||
DB 180 AATGCGCTTCGAAACACGAGTGAAGTGAAGCTTTCTTGCCATTTTGTGATGAG 239
QY 97 CysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPhe 116
|||||
DB 240 TGTGCGACCTTAAAGGTTTGGATTCATCTTGTGCATGACATGATCAATTCACATTT 299
QY 117 GluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
|||||
DB 300 GAGTCTGAGGTACCAATGATTAACCAAGCAGTGAAGCTCTGTAATAACTGAC----- 353
QY 137 IlePheGluGluGluLysSerAspAsp-----AspLysPheGluProPheSerLeu 153
|||||
DB 354 ATATTGAATCAGAAAGATTTGCTGATGAGATTAATTCACATCAACCAACCTTTCTTTC 413
QY 154 CysSerLysProArgLysArgArgGlnArgLysGlyLysArgAsn 168
|||||
DB 414 TGTTCAGATCTCGAAACGTAAGACAGAGGCTCCGTTCAAAAT 458

RESULT 13

LOCUS BG314044

DEFINITION WHE2467_D04_G07Z5 Trifolium monocoecum early reproductive apex cDNA
405 bp mRNA linear EST 16-APR-2001


```

OY 277 PrometAlaLeuGluValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 505 CCAATGCACTGGAGCAAGTGTCTCGACCGTGTAGTGAAGATGATGATGATGAT 446
OY 297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAspGlu 316
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 445 ATGTGACTTGGAGATGAGAAGATGCTGTGATGATTTGTTGATGATTCAGAAATGAA 386
OY 317 LysGlnPheMetHisLeuTrpAsnSerPheValArgGlyGlnArgValIleAlaAspGly 336
    ||| ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 385 AAACCTATTATGATATGATGATGATTCATTTCTCAAAACAAAGC----- 341
OY 337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTrpGlyGluLeuHisArgTyr 356
DB 341 ----- 341
OY 357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuVal 376
    ||||| ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 340 -----GTTGGCGTTTCTTCATGATTAACCTTTGGAACCATTAATTTTA 296
OY 377 AsperAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 295 GATGCCCGCACTATGAAACACATGCAATACAGTCTTCAAAATTTTACAAGAAAGAAC 239

RESULT 15
BE203328 612 bp mRNA linear EST 07-SEP-2000
LOCUS BE203328 KVI1 Medicago truncatula cDNA clone pKV1-5M13, mRNA
DEFINITION EST403350 KVI1 Medicago truncatula cDNA clone pKV1-5M13, mRNA
ACCESSION BE203328
VERSION BE203328.1 GI:8746599
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
            Eudicotyledons; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
            Medicago.
REFERENCE 1 (bases 1 to 612)
AUTHORS Vandenbosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L.,
        Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
        Fraser,C.M.
        ESTs from roots of Medicago truncatula 24 hours after inoculation
        with Sinorhizobium meliloti
        unpublished (1999)
JOURNAL Contact: Vandenbosch K
COMMENT Department of Plant Biology
        University of Minnesota
        220 Bloss Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
        Tel: 612 624 2755
        Fax: 612 625 1738
        Email: kvandenbosch.umn.edu
        Texas A&M University: T263100e
        TIGR sequence name: MTIAM797K
        More information is available at:
        http://chrysis.tamu.edu/medicago
        Seq primer: SKmod (CTA gaa cta gtc gat cc).
        Location/Qualifiers
            1..612
                /organism="Medicago truncatula"
                /cultivar="genotype A17"
                /db_xref="taxon:3880"
                /clone="pKV1-5M13"
                /clone_1kb="KVI1"
                /tissue_type="Seedling roots"
                /dev_stage="24 hours post-inoculation with Sinorhizobium
                meliloti"
                /lab_host="E. coli strain XL04R"
                /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                was directionally ligated into the Unizap XR vector from
                Stratagene and packaged using Gigapack III Gold packaging

```

extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL04R cells."

BASE COUNT 200 a 88 c 132 g 192 t

ORIGIN

Alignment Scores:

Score:	7.23e-41	Length:	612
Percent Similarity:	44.250	Matches:	79
Best Local Similarity:	81.158	Conservative:	20
Query Match:	64.758	Mismatches:	18
DB:	18.618	Indels:	5
	Gaps:		1

US-09-890-220-2 (1-445) x BE203328 (1-612)

```

OY 293 ValAlaAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspVal 312
    ||||| ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 3 GTTGATGATGAGTGTCTAACCCTGGAAGATGAAAGATGATGATTTGTGATGTT 62
OY 313 AsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgGlyGlnArgVal 332
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 63 TCCAAAGATGAAAGACACCTCATGCACTCTGGAACCTTTATAGAAACAAAGGCTA 122
OY 333 IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTrpGlyGlu 352
    ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 123 CTGGCTGATGTCATGCTTCCTCGGCTGTGTGAGCATTTTCCAAACATTTATCCAAAGAC 182
OY 353 LeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 183 CTGAATTCATCCCGAACCTTATTTGGCTTGAGATTAATCATGCTCAAACTTGGAAAT 242
OY 373 HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392
    ||||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 243 CATGGCTTCTTGATGATCAACCAATGAAACACTCAGATATATTTGAAAGCAACCA 302
OY 393 AsnSerSerAspThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSer 412
    ||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
DB 303 AACGGGGATCAGACAGTGCACAAAT-----TGAAGACCATCTGTCTCA 347
OY 413 AsnThr 414
    ::|||
DB 348 TCCACA 353

```

Search completed: June 20, 2003, 01:23:27
Job time : 1937 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:23:13 ; Search time 23 Seconds

(without alignments)
802.477 Million cell updates/sec

Title: US-09-890-220-2

Sequence: 2378 1 MCRONCRAKSSPEVISTDE.....INNNNNVDKNNSRDVKYIK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138.5	5.8	989	1	PRP3_DICDI
2	132	5.6	749	1	MAD1_YEAST
3	127.5	5.4	448	1	AAC2_DICDI
4	126.5	5.3	1584	1	KYK1_DICDI
5	123.5	5.2	1858	1	P3K2_DICDI
6	120	5.0	1000	1	S155_YEAST
7	120	5.0	1195	1	YK76_YEAST
8	119	5.0	1094	1	YB00_YEAST
9	117.5	4.9	988	1	OMB_DROME
10	117	4.9	2175	1	HMCU_DROME
11	114	4.8	964	1	YIN0_YEAST
12	112.5	4.7	2278	1	FAB1_YEAST
13	112	4.7	1314	1	SM11_YEAST
14	111.5	4.7	735	1	C1GB_DICDI
15	111.5	4.7	858	1	CYAG_DICDI
16	108	4.5	1407	1	CYAA_DICDI
17	107.5	4.5	833	1	HSE_YEAST
18	107.5	4.5	2339	1	PC11_PLAFA
19	106.5	4.5	1585	1	P3K3_DICDI
20	106	4.5	694	1	HAAL_YEAST
21	106	4.5	779	1	SRP_DROME
22	105.5	4.4	558	1	YCX1_ASTLO
23	105	4.4	396	1	RMAR_YEAST
24	105	4.4	555	1	POST_XENLA
25	105	4.4	732	1	KMHB_DICDI
26	104.5	4.4	666	1	YEA7_YEAST
27	104.5	4.4	935	1	COPE_YEAST
28	104.5	4.4	1419	1	MDR_PLAFA
29	104	4.3	871	1	SC10_YEAST
30	103	4.3	432	1	AP2_ARATH
31	103	4.3	490	1	CAR3_DICDI
32	103	4.3	1085	1	CUT7_SCHRO
33	102.5	4.3	481	1	BIND_STRFO

34	102.5	4.3	1076	1	RP08_ASTIO
35	102.5	4.3	1570	1	P3K1_DICDI
36	102.5	4.3	2733	1	RRPB_CWAS
37	101.5	4.3	647	1	KNRL_DROME
38	101.5	4.3	946	1	FTSR_CAMJE
39	101.5	4.3	1341	1	YL78_YEAST
40	101.5	4.3	2731	1	RRPB_CWMT
41	101	4.2	534	1	GCR2_YEAST
42	101	4.2	537	1	ARP_PLAFA
43	101	4.2	1131	1	YAB9_YEAST
44	100.5	4.2	1188	1	OSH1_YEAST
45	100	4.2	284	1	SMX5_SCHMA

ALIGNMENTS

RESULT 1	ID	PRP3_DICDI	STANDARD:	PRT:	989 AA.
AC	P54637				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Protein-tyrosine phosphatase 3 (EC 3.1.3.48) (Protein-tyrosine-phosphatase phosphohydrolase 3).				
DE	(PRP1 OR PRP3) AND (PTPC2 OR PTP3).				
GN	Dictyostellum discoideum (Slime mold).				
OS	Dictyostellum discoideum (Slime mold).				
OC	Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.				
OX	NCBI_TaxID=44689;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=AX3.				
RX	MEDLINE=96189126; PubMed=8628311;				
RA	Gamper M., Howard P.K., Hunter T., Firtel R.A.;				
RT	"Multiple roles of the novel protein tyrosine phosphatase PRP3 during Dictyostellum growth and development."				
RL	Mol. Cell. Biol. 16:2431-2444(1996).				
CC	-1- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 kDa (P130).				
CC	-1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O -> protein tyrosine + phosphate.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	-1- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.				
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH AND DEVELOPMENT.				
CC	-1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: U38197; AAC47041.1; -				
DR	HSSP: 006124; 2SRP.				
DR	DictyDB: DD01111; PEPCL.				
DR	DictyDB: DD07777; PEPCL.				
DR	InterPro: IPR000387; Tyr_Pp.				
DR	InterPro: IPR000242; Tyr_Pp.				
DR	Pfam: PF00102; Y-phosphatase; 1.				
DR	PRINTS: PR00700; PRTYPPHATASE.				
DR	SMART: SM00194; PTPc; 1.				
DR	PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.				
DR	PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.				
DR	PROSITE: PS50055; TYR_PHOSPHATASE_Pp; 1.				
KW	Hydrolase.				
FT	ACT_SITE	649	649		
FT	DOMAIN	460	716		BY SIMILARITY. PROTEIN-TYROSINE PHOSPHATASE.
FT	DOMAIN	64	71		POLY-ASN.
FT	DOMAIN	109	118		POLY-ASN.


```

ID AAC2.DICDI STANDARD; PRT; 448 AA.
AC P14196;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE AAC-rich mRNA clone AAC11 protein (Fragment).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
ON NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90066348; PubMed=2511421;
RA Shaw D.R., Richter H., Giorda R., Omachi T., Ennis H.L.;
RT "Nucleotide sequences of Dictyostelium discoideum developmentally
RT regulated cDNAs rich in (AAC) imply proteins that contain clusters of
RT asparagine, glutamine, or threonine."
RL Mol. Gen. Genet. 218:453-459(1989).
CC -1- DEVELOPMENTAL STAGE: THE CONCENTRATION OF AAC-RICH MRNAS IS LOW
CC IN DORMANT SPORES AND GROWING CELLS, BUT INCREASES DURING
CC SPORE-GERMINATION AND MULTICELLULAR DEVELOPMENT.
CC -1- METEOLANEOUS: SEVERAL PROTEINS DERIVE FROM AAC-RICH MRNA, WHICH,
CC DUE TO A FRAMESHIFT ALSO HAVE ACA AND CAA CODONS AND THUS ARE
CC ASN-, THR- OR GLN-RICH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16522; CAA34529.1; -.
DR PIR; S05355; S05355.
DR DictyDb; DD05006; -.
DR InterPro; IPR000637; AT_hook.
DR InterPro; IPR000116; Highmobility-IT.
DR Pfam; PF02178; AT_hook; 4.
DR ProDom; PD005593; Highmobility-IT; 1.
DR SMART; SM00384; AT_hook; 4.
KW Repeat.
FT NON_TER 1 1
FT DOMAIN 5 25 GLN-RICH.
FT DOMAIN 115 144 ASN-RICH.
FT DOMAIN 268 384 ASN-RICH.
SQ SEQUENCE 448 AA; 48636 MW; D8FDB0DD910D99817 CRC64;

Query Match 5.48; Score 127.5; DB 1; Length 448;
Best Local Similarity 35.78; Pred. No. 0.016;
Matches 25; Conservative 20; Mismatches 24; Indels 1; Gaps 1;

QY 372 NHGLVDAATINCNCWTILENCSSDDTGTNNNSVDRSDSNTNNNTIVDHPNDINKNN 431
   1: ::::||||| | ::::||||| : ::::|||||
Db 273 NNNNNNNNSNNNNNNNNNNNNNNNNNNSSNNNNNN-NNSTNNNTNNNTNNNTNNNNNN 331
   1: ::::|||||
QY 432 VDNKDNNSRD 441
   1: :|||:
Db 332 INNNTNNNTNN 341
   1: :|||:

RESULT 4
ID KYK1.DICDI STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysin A (EC 2.7.1.112) (Tyrosine-
DE protein kinase 1).
DE PYKA OR SPIA OR DPYK1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
ON NCBI_TaxID=44689;

```

RP	SEQUENCE FROM N.A.	[1]
RC	STRAIN-JH10;	
RA	MEDLINE=97053827; PubMed=8898241;	
RA	Nuckolls G.H., Oshero N., Loomis W.F., Spudich J.A.;	
RT	"type Dictyostellum dual-specificity kinase splA is essential for spore differentiation".;	
RL	Development 122:3295-3305(1996).	
RN	[2]	
RP	SEQUENCE OF 1248-1584 FROM N.A.	
RX	MEDLINE=90287147; PubMed=1972546;	
RA	Tan J.L., Spudich J.A.;	
RT	"Developmentally regulated protein-tyrosine kinase genes in Dictyostellum discoideum".;	
RL	Mol. Cell. Biol. 10:3578-3583(1990).	
CC	-1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.	
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.	
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK DURING THE MOUND STAGE OF MORPHOGENESIS.	
CC	-1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.	
CC	-1- SIMILARITY: CONTAINS 1 SAM DOMAIN.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; U32174; AAB41125.1; -;	
DR	EMBL; M33785; AAA33202.1; -;	
DR	PIR; A35670; A35670.	
DR	Dictydb; DD03010; PYKA.	
DR	InterPro; IPR000719; Euk_pkinase.	
DR	InterPro; IPR001660; SAM.	
DR	InterPro; IPR003878; SPRY_domain.	
DR	InterPro; IPR003877; SPRY_receptor.	
DR	InterPro; IPR004040; TRY_kinase.	
DR	InterPro; IPR001245; Tyr_pkinase.	
DR	Pfam; PF00069; pkinase; 1.	
DR	Pfam; PF00536; SAM; 1.	
DR	Pfam; PF00622; SPRY; 3.	
DR	ProDom; PD000001; Euk_pkinase; 1.	
DR	SMART; SM00454; SAM; 1.	
DR	SMART; SM00449; SPRY; 3.	
DR	SMART; SM00221; STYK; 1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PS05001; PROTEIN_KINASE_DOM; 1.	
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	
DR	PROSITE; PS50105; SAM_DOMAIN; 1.	
KW	Transferase; Tyrosine-protein kinase; ATP-binding; phosphorylation.	
FT	DOMAIN 908 972 SAM.	
FT	DOMAIN 403 420 POLY-ASN.	
FT	DOMAIN 428 435 POLY-THR.	
FT	DOMAIN 449 480 POLY-ASN.	
FT	DOMAIN 483 491 POLY-ASN.	
FT	DOMAIN 494 508 POLY-ASN.	
FT	DOMAIN 512 532 POLY-ASN.	
FT	DOMAIN 596 600 POLY-ASN.	
FT	DOMAIN 808 811 POLY-PHE.	
FT	DOMAIN 1026 1029 POLY-SER.	
FT	DOMAIN 1195 1210 POLY-ASN.	
FT	DOMAIN 1215 1220 POLY-GLN.	
FT	DOMAIN 1224 1233 POLY-GLN.	
FT	DOMAIN 1266 1274 POLY-PRO.	
FT	DOMAIN 1289 1561 PROTEIN_KINASE.	
FT	NP_BIND 1295 1303 ATP (BY SIMILARITY).	
FT	NP_BIND 1316 1316 ATP (BY SIMILARITY).	
FT	ACT_SITE 1417 1417 BY SIMILARITY.	
FT	CONFLICT 1248 1248 D -> R (IN REF. 2).	
FT	CONFLICT 1435 1435 V -> L (IN REF. 2).	

SEQUENCE 1584 AA: 174304 MW: 5D1589458DBE01E3 CRC64;

Query Match 5.3%; Score 126.5; DB 1; Length 1584;
Best Local Similarity 22.2%; Pred. No. 0.099; Mismatches 100; Indels 133; Gaps 15;
Matches 77; Conservative 37; Mismatches 100; Indels 133; Gaps 15;

DB 149 EEPFSCSKPRKRROGGRNMTRLKVCFLPLDPSLNGTENGITLLDNGRGGLGYPEAT 208
DB 210 EEEFECRSEKKNYH-GKHVYVRNTAFLLPDSPTDITG---GVATQPFEGECY---- 261
DB 209 ELAOGFEMTSNIPPAHSSLDAGAKVILLTSEAVVPATKRLKLSERSEARSHLLQKQ 268
DB 262 -----FE-----VIIDQDKGOLSTIGLANLEPFP-----YHGVMPRS 295
DB 269 F-YHSHRYQPMALQVMSDRSEDEVDVADFEEDROMLDDFVNNKDEKQFMHLMNSFV 327
DB 296 YGYH-----NDDGKRFRMRRE---PGVNEGES-----YGSY 324
DB 328 RKQRYADGHISMACE-----AFSREYKELHRYSSLFMCWRLP----- 366
DB 325 KKQDIYGGC-LSEFSREIFETKNGMYLGTAFSANY-----GVFQPSVAFNPGISIT 375
DB 367 -----LKLWHLGVDSATI-----NNCNTILENCNSSDGT 398
DB 376 GVFGPPFKFSOVTLMKLN---VNSTSLVPGNKNNNNNNNNNNNNNNNNNNNNNNN 432
DB 399 TT-----NNNSVDRPSDSMTNNNTVDHPNDINKNNVNDKDNNS 439
DB 433 TSTSPSTNNNEDISSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 479

RESULT 5

P3K2_DICDI STANDARD; PRT; 1858 AA.

AC P34674; 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Phosphatidylinositol 3-kinase 2 (EC 2.7.1.137) (PI3-kinase)
DE (Ptdins-3-kinase) (PI3K).
GN PI3K OR PIK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takagawa K., Emr S.D., Firtel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
discoideum: Biological roles of putative mammalian p110 and yeast
Vps34p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC 1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-ID-myo-inositol = ADP +
1-phosphatidy1-ID-myo-inositol 3-phosphate.
CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U23477; AAA85722.1;
DR Dictydb: DD01100; PI3K.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR000341; PI3K_Cas_Bind.
DR InterPro: IPR01263; PI3K.
DR InterPro: IPR00403; PI3_P14_kinase.
DR Pfam: PF00454; PI3_P14_kinase; 1.
DR Pfam: PF00613; PI3Ka; 1.

DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00794; PI3K_C2; 1.
DR SMART: SM00142; PI3K_C2; 1.
DR SMART: SM00144; PI3K_C2; 1.
DR SMART: SM00145; PI3Ka; 1.
DR SMART: SM00146; PI3K; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS00917; PI3_4_KINASE_3; 1.
DR PROSITE: PS00918; PI3_4_KINASE_4; 1.
KW Transferase; Kinase; Multigene family.
FT DOMAIN 34 40 POLY-GLY.
FT DOMAIN 166 172 POLY-SER.
FT DOMAIN 185 226 POLY-ASN.
FT DOMAIN 227 235 POLY-THR.
FT DOMAIN 246 253 POLY-SER.
FT DOMAIN 259 268 POLY-ASN.
FT DOMAIN 294 303 POLY-THR.
FT DOMAIN 361 364 POLY-SER.
FT DOMAIN 369 384 POLY-SER.
FT DOMAIN 425 429 POLY-GLN.
FT DOMAIN 439 444 POLY-THR.
FT DOMAIN 445 454 POLY-SER.
FT DOMAIN 562 570 POLY-GLY.
FT DOMAIN 715 727 POLY-THR.
FT DOMAIN 982 990 POLY-GLN.
FT DOMAIN 1015 1049 POLY-ASN.
FT DOMAIN 1598 1858 POLY-ASN.
SQ SEQUENCE 1858 AA: 203945 MW: A6C033D4CDEA03 CRC64;

Query Match 5.2%; Score 123.5; DB 1; Length 1858;
Best Local Similarity 27.5%; Pred. No. 0.21; Mismatches 38; Conservative 25; Mismatches 42; Indels 33; Gaps 5;

DB 312 VNDEKQFMH-LMNSFVRKQRYADGHISMACEAFSREYKELHRYSSLFMCWRLFLKL 370
DB 113 IEKEKEIINTISNSGVTFYMT-----ALEIIDSTINTPLNRSR----- 152
DB 371 WNLGLVDSATINNCNTILENCNS-----SDTTTNNNSVDRPSDNTNNNTVDHP 423
DB 153 ---GSIGSKPI-CNNLTSSSSSSSTTATTPPTTSSNNNNNNNNNNNNNNNNNNNNNN 207
DB 424 NDIINKNNVNDKDNNSRD 441
DB 208 NNN 225

RESULT 6

S153_YEAST STANDARD; PRT; 1000 AA.

AC P43612; 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S154-associating protein SAP155.
GN SAP155 OR YFR040W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Andre K.T., "The SAP, a new family of proteins, associate and function positively
with the S154 phosphatase."
RL Mol. Cell. Biol. 16:2744-2755(1996).
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=5288c / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shidrata T., Ozawa M.,
RA Sasamura S.-I., Sasamura M., Tsuchiya Y., Seda E., Yokoyama K.,

```

05 Saccharomyces cerevisiae (Baker's yeast).
0C Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
RN NCBI_TaxID=4932;
[1]
RP SEQUENCE OF 1-843 FROM N.A.
RA Baladron V., Ballesta J.P.G., Bou G., del Rey F., Esteban P.F.,
RA Garcia-Cantalejo J.M., Garcia-Ramirez J.J., Gonzalez A., Jimenez A.,
RA Revuelta J.L., Santos M.A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 737-1195 FROM N.A.
RA Gallion L., Dujon B.;
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO YEAST YII151C. ALSO TO THE N-TERMINAL OF
CC YEAST ALDEHYDE DEHYDROGENASE, MITOCHONRIAL 1 (ALD1).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC CC
CC DR EMBL; Z28321; CAA82176.1; -.
CC DR PIR; S38174; S38174.
CC DR SGD; S0001804; YKR096W.
CC KW Hypothetical protein; ATP-binding.
CC FT DOMAIN 210 250 ASN-RICH.
FT NP_BIND 465 472 ATP (POTENTIAL).
SQ SEQUENCE 1195 AA; 137490 MW; D7251DD6B523622 CRC64;

Query Match          5.0%; Score 120; DB 1; Length 1195;
Best Local Similarity 21.4%; Pred. No. 0.22;
Matches    62; Conservative   37; Mismatches    99; Indels     92; Gaps    10;

QY      166 RNNTRRLKVCFLPLDSPSLTNGTENGITLLNDGNRGIGYEATELAGQ-----213
DB      42 RNTTANLSLSFOGYLPNPSKKRONSNYSYVDIINSSKRISRPRFSDIEGKNDDHTYPERTT 101
QY      214 -FEWTSNPPIPAIAHSSLDAGAKVILLTAAYVPARKTKRL-SAESEARSHELLOKRPFYH 271
DB      102 VKESKPNNSPPRYVSSS---KRALKRENSVGITQSALSISKSPENSGSIJAHNR--WSP 154
QY      272 SHRQPMALDEQWMSDRosedEVDVDVADEFEDQLMDPEVDVYNKDEKQFMHLMNSFYVRQR 331
DB      155 ENMKIKPLAVSO-----NSLA-----FVDAGSDEQSKSELVYGCFORKSN 192
QY      332 VIADGHISMACEAFSRFEKELHRRSYLSLFMCWRFLIKLMNHGLVDSATINNCTILENC 391
DB      193 -----NSGEINDKD--NSA 204
QY      392 RNSSDTTTNNNNNSYDRSPDSNTNNNNIVDPHPDINKKNVNDKDNNSRD 441
DB      205 RDQDFNNGNNNNNNNNNSSNNNDNNNNNDNNNNNNNNNSNSRDN--NNSSD 253

RESULT 8
ID_YB00_YEAST STANDARD; PRT; 1094 AA.
AC P38114;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative 126.9 kDa transcriptional regulatory protein in YSM1-RIB7
DE intergenic region.
GN YBR150C OR YBR1133.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
RN [1]
```

RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RA Ntlian K.-D., Koeltter P., Rose M., Becker J., Grey M., Li Z.,
RA Niedmann E., Schenk-Groening R., Serrus J., Weimer E.,
RA Wolter R., Brendel M., Bauer J., Brann H., Derr K., Diesner S.,
RA Gruenbein R., Heddes J., Kriesau P., Korol S., Krems B., Frott M.,
RA Siegers K., Bauer A., Boles E., Miosga T.,
RA Schaaff-Gershenchleger I., Zimmermann F.K.,
RA Submitted AUG 1994 to the EMBL/GenBank/DBS databases.
CC 1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC 1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC CLUSTER DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)
CC or_send_an_email_to_license@sib-sib.ch).
CC -----
CC EMBL, Z36019; CA85108.1; -
CC PIR, S46021; S46021.
CC HSSP, P08657; 1CID.
CC TRANSFAC, T03677; -
CC SGD, S0000354; YBR150C.
CC InterPro, IPR001138; Fungal-Trn.
CC Pfam, PF00172; Zn-clus; 1.
CC SMART, SM00066; GAL4; 1.
CC PROSITE, PS00463; ZN2_CYG_FUNGAL_1; 1.
CC PROSITE, PS50048; ZN2_CYG_FUNGAL_2; 1.
CC Hypothetical protein: Transcription regulation, DNA-binding;
CC Nuclear protein; Zinc; Metal-binding.
CC DNA_BIND 107 137
CC FT SEQUENCE 1094 AA; 126902 MW; 3C0BF8F5679D14F CRC64;
SQ

Query Match 5.08; Score 119; DB 1; Length 1094;
Best Local Similarity 27.58; Pred. NO. 0.23;
Matches 44; Conservative 26; Mismatches 44; Indels 46; Gaps 9;

QY 303 ROMIDDFYDVNKKDEKOFMHLNLSYVKRQVYADGHISNAC-----AFSREYE--KE 352
DB 859 RESKDQWELKK-----TSFINNEKMAAFG--SEVTDQEKMEVSFIMFEDITQD 907
QY 353 LH--RYSDFEWCWRLFLIKINHGILVDSATINNCNTILLENCRNSDDTTTNNNS-VDRP 409
DB 908 LNFQFSS-----IPKLEWENTLEEG-----EEYHSHNGTVDNNEETGGADDT 950
QY 410 SDSNTNNNTVDHPNDI-----NNKNVNDKNNNSRDKVI 444
DB 951 DDNNNNNNNNNNKNGNNSSTINNNNNYSNNSNNDNDNNI 990

RESULT 9
OMB_DROME STANDARD: PRT: 988 AA.
AC Q24432; Q27917; Q5W4K5;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Optomotor-blind protein (lethal(1)optomotor-blind) (L(1)omb) (Bifid
DE protein).
GN BI OR OMB OR CG3578.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R; TISSUE-Embryo;
RC MEDLINE=92159016; PubMed=1741374;
RA Pfingfelder G.O., Roth H., Poock B., Kerscher S., Schwarz H.,

RA Jonschok B., Helsenberg M. ;
 RT "The lethal(1)optomotor-blind gene of *Drosophila melanogaster* is a
 RT major organizer of optic lobe development: Isolation and
 RT characterization of the gene".
 RL Proc. Natl. Acad. Sci. U.S.A. 69:1199-1203(1992).
 RN [2].
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132.
 RA Adams M.D., Gelinkter S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amenthalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bollock S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotliker P.,
 RA Butts K.C., Busam D.A., Butler H., Cadenhead E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Deodson K., Doup L.E., Downes M., Dugan-Necha S., Dunkov B.C., Dunn P.,
 RA Fowler K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,
 RA Jalali M., Kautus F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Ralp D., Lai Z.,
 RA Lahn X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon R., Nussken D.R., Paclio J.M.,
 RA Palazzolo M., Platan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E.R., Spradling A.C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Svirskas R., Tector C., Threlton M., Strong R., Sun E.,
 RA Wang Z.-Y., Wassarman D.A., Welnslock G.M., Weissensbach J.,
 RA Williams S.M., Woodard T., Wolley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhu Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gbbs R.A., Myers E.W., Rubin G.M., Venter J.C. ;
 RA "The genome sequence of *Drosophila melanogaster*".
 RL Science 287:2185-2195(2000).
 RN [3].
 RP SEQUENCE OF 1-447 FROM N.A., AND MUTATIONAL ANALYSIS.
 RC TISSUE=Larva;
 RX MEDLINE=93261414; PubMed=8492800;
 RA Poock B., Balles J., Pfeufferfeld G.O. ;
 RT "transcript identification in the optomotor-blind locus of *Drosophila*
 RT melanogaster by intragenomic recombination mapping and PCR-aided
 RT sequence analysis of lethal point mutations".
 RL Mol. Gen. Genet. 238:325-332(1993).
 CC -1- FUNCTION: ESSENTIAL PROTEIN THAT MAY FUNCTION AS A TRANSCRIPTION
 CC REGULATOR. FLIES WITH L(1)OMB MUTATIONS SHOW SEVERE MALDEVELOPMENT
 CC OF THE OPTIC LOBES. REDUCTION IN WING SIZE AND AN INCREASED
 CC ABDOMINAL PIGMENTATION. THEY DIE DURING THE Pupal STAGE.
 CC -1- SUBCELLULAR LOCATION: Nucleus (tentative).
 CC -1- TISSUE SPECIFICITY: IN THIRD-INSTAR LARVAE IT IS FOUND IN THE
 CC BRAIN REGION THAT WILL DEVELOP INTO OPTIC LOBES AND MORE WEAKLY IN
 CC THE THORACIC PART OF THE VENTRAL GANGLION.
 CC -1- DEVELOPMENTAL STAGE: THE PEAK PERIODS OF EXPRESSION ARE: MID-
 CC EMBRYOGENESIS, THE SECOND DAY OF Pupal DEVELOPMENT AND IN THE
 CC ADULT.
 CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M81796; AAA8736.1; -
 DR EMBL; AE003431; AAF45946.1; -
 DR EMBL; S61732; AAB26697.1; -
 DR EMBL; S61727; AAB26697.1; JOINED.
 DR EMBL; S61729; AAB26697.1; JOINED.
 DR EMBL; S61744; AAB26699.1; -
 DR EMBL; S61743; AAB26699.1; JOINED.
 DR EMBL; S61955; AAB26699.1; JOINED.
 DR HSSP; P24781; 1XBR.
 DR FLYBASE; FBgn0000179; b1.
 DR InterPro; IPR001699; TF_T-box.
 DR Pfam; PF00907; T-box; 1.
 DR PRINTS; PR00937; TBOX.
 DR SMART; SM00425; TBOX; 1.
 DR PROSITE; PS01283; TBOX_1; 1.
 DR PROSITE; PS01264; TBOX_2; 1.
 DR PROSITE; PS50252; TBOX_3; 1.
 KM DNA-binding; Nuclear protein; Transcription regulation.
 FT DOMAIN 51 140
 FT DOMAIN 104 107 ASN-RICH.
 FT DOMAIN 179 184 POLY-THR.
 FT DOMAIN 229 236 POLY-SER.
 FT DOMAIN 238 244 POLY-GLN.
 FT DNA_BIND 332 513 POLY-PRO.
 FT DOMAIN 574 577 T-BOX.
 FT DOMAIN 607 692 POLY-ASP.
 FT DOMAIN 823 831 ALA-RICH.
 FT DOMAIN 910 916 POLY-GLY.
 FT DOMAIN 926 966 POLY-ALA.
 FT CONFLICT 10 10 GLN/HIS-RICH.
 FT CONFLICT 216 216 F->L (IN REF. 1 AND 3).
 FT CONFLICT 511 511 A->P (IN REF. 1 AND 3).
 FT CONFLICT 823 823 F->L (IN REF. 1).
 FT CONFLICT 976 988 MISSING (IN REF. 1).
 SQ SEQUENCE 988 AA; 103992 MW; 032B7A4471743FC9 CMC64;

Query Match 4.9%; Score 117.5; DB 1; Length 988;
 Best Local Similarity 35.2%; Pred. No. 0.26;
 Matches 32; Conservative 11; Mismatches 23; Indels 25; Gaps 3;

QY 372 NHGLVDSATNNCNLENCN-----SSDTTNNNSVDRSD 411
 DB 59 NSGNNSNSNNNTNTNTNVLVAVSPGGAQLSPSHSSNTTTSNTN-----SS 114
 QY 412 SNTNNNNIVDHPNDINNNKNDKDNNSRD 442
 DB 115 SNNNNNST-HNNNNHTNNNNNNNNNTSOK 144

RESULT 10
 HMCU_DROME STANDARD; PRT; 2175 AA.
 AC P10180; 09N306; -
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homeobox protein cut.
 GN CT OR CG1387.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88232956; PubMed=2897632;
 RA Blochlinger K., Bodmer R., Jack J., Jan L.Y., Jan Y.N.;
 RT "Primary structure and expression of a product from cut, a locus
 RT involved in specifying sensory organ identity in Drosophila.";

RL Nature 333:629-635(1988).
 RP [2]
 RC SEQUENCE FROM N.A.
 RX STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikler S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Chang M., Flier B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abriell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Teclor C., Turner C., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: REGULATOR OF CELL FATE DECISIONS IN MULTIPLE LINEAGES.
 CC SPECIFICALLY, FUNCTIONS AS A DETERMINATION FACTOR THAT SPECIFICS
 CC SENSOR ORGAN IDENTITY IN PRECURSOR CELLS. PROBABLY ALSO INVOLVED
 CC IN CELL TYPE SPECIFICATION OF MALPIGHIAN TUBULES. IN ABSENCE OF
 CC CUT GENE EXTERNAL SENSOR ORGANS ARE TRANSFORMED INTO CHORDOTONAL
 CC ORGANS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: DETECTED IN MANY CELLS IN THE CENTRAL NERVOUS
 CC SYSTEM, ALL EXTERNAL SENSOR ORGANS, SOME PERIPHERAL NEURONS, AND
 CC IN THE NON-NEURAL CELLS OF THE SPIRACLES AND THE MALPIGHIAN
 CC TUBULES.
 CC -1- DEVELOPMENTAL STAGE: CELL-SPECIFIC PATTERN OF EXPRESSION. BROADLY
 CC EXPRESSED DURING EMBRYONIC DEVELOPMENT.
 CC -1- DOMAIN: ASN AT POSITION 47 OF THE HOMEBOX MAY PARTICIPATE IN
 CC REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
 CC HETERO-DIMERIZATION.
 CC -1- SIMILARITY: CONTAINS 3 CUT DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE CUT HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).


```

DR TRANSFAC; T02004;
DR FlyBase; FBgn0004198; ct.
DR InterPro; IPR000047; HTH_repressr.
DR InterPro; IPR003350; Homeo CUT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00031; HTHREPRESSR.
DR PRODOM; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR Transcription regulation; Homeobox; DNA-binding; Coiled coil.
KW Developmental protein; Nuclear protein; Repeat; Coiled coil.
FT DOMAIN 265 343
FT DNAS_BIND 433 499
FT DNAS_BIND 798 964
FT DNAS_BIND 1056 1161
FT DNAS_BIND 1329 1417
FT DNAS_BIND 1463 1522
FT DNAS_BIND 1608 1695
FT DNAS_BIND 1745 1804
FT DNAS_BIND 194 210
FT DNAS_BIND 235 243
FT DNAS_BIND 271 293
FT DNAS_BIND 384 428
FT DNAS_BIND 547 554
FT DNAS_BIND 574 584
FT DNAS_BIND 616 630
FT DNAS_BIND 665 699
FT DNAS_BIND 2004 2014
FT DNAS_BIND 2071 2077
FT DNAS_BIND 2124 2136
FT DNAS_BIND 2175 AA; 233628 MW; 08BF80C4861BD0AB CRC64;
SQ SEQUENCE

Query Match 4.9%; Score 117; DB 1; Length 2175;
Best Local Similarity 23.1%; Pred. No. 0.8;
Matches 58; Conservative 32; Mismatches 107; Indels 54; Gaps 8;

```

```

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-5288c / Ab972;
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentes S., Hamlyn N., Horsnell T.S., Hunt S., Jagsels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINDING
CC CLUSTER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z38059; CA86148.1;
CC PIR; S48404; S48404.
CC HSP; P08657; IGLD.
CC SGD; S0001392; YJL130W.
CC InterPro; IPR001138; Fung1_Tfn.
CC Pfam; PF00172; Zn_c1us; 1.
CC PRINTS; PR00054; FUNGALINCS.
CC SMART; SM00066; GAL4; 1.
DR PROSITE; PS00463; ZN2_CYS6_FUNGAL_1; 1.
DR PROSITE; PS50048; ZN2_CYS6_FUNGAL_2; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Nuclear protein; Zinc; Metal-binding.
FT DNAS_BIND 21 47
FT DNAS_BIND 811 896
FT DNAS_BIND 964 AA; 108780 MW; AD5AE5E48022CC CRC64;
SQ SEQUENCE

Query Match 4.8%; Score 114; DB 1; Length 964;
Best Local Similarity 40.4%; Pred. No. 0.47;
Matches 23; Conservative 11; Mismatches 21; Indels 2; Gaps 1;

```



```

RA STRAIN-S288C / AB972:
RX MEDLINE-95400292: Pubmed-7670463:
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.,
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae ";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SIMILARITY TO TCPI/CCP FAMILY.
RX MEDLINE-95314774: Pubmed-7794526:
RA Waldmann T., Lupas A., Kellermann J., Peters J., Baumeister W.;
RT "Primary structure of the thymosome from Thermoplasma acidophilum.";
RL Biol. Chem. Hoppe-Seyler 376:119-126(1995).
CC -1- FUNCTION: CATALYSES THE PHOSPHORYLATION OF PHOSPHATIDYLINOSITOL-4-
CC PHOSPHATE ON THE FIFTH HYDROXYL OF THE MYO-INOSITOL RING, TO FORM
CC PHOSPHATIDYLINOSITOL-4,5-BIPHOSPHATE. REQUIRED FOR ENDOCYTIC-
CC VACUOLAR PATHWAY AND NUCLEAR MIGRATION. THE PRODUCT OF THE
CC REACTION IT CATALYSES FUNCTIONS AS AN IMPORTANT REGULATOR OF
CC VACUOLE HOMEOSTASIS PERHAPS BY CONTROLLING MEMBRANE FLUX TO AND/OR
CC FROM THE VACUOLE.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-ID-myo-inositol 4-
CC monophosphate -> ADP + 1-phosphatidy1-ID-myo-inositol 4,5-
CC biphosphate.
CC -1- SUBCELLULAR LOCATION: VACUOLE-ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PDIINS(4)P-5-KINASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U01017; AAA81360.1; -;
DR EMBL: D50617; BAA09258.1; -;
DR SGD: S0001915; FABI.
DR InterPro: IPR002498; PIP5K.
DR InterPro: IPR000306; Znf_FYVE.
DR Pfam: PF01363; FYVE; 1.
DR Pfam: PF01504; PIP5K; 1.
DR SMART: SM00064; FYVE; 1.
DR SMART: SM00330; PIPK; 1.
DR PROSITE: PS50178; ZF_FYVE; 1.
KW Transferase; Kinase; zinc-finger; zinc.
FT ZN_FING 240 299 FYVE-TYPE.
FT DOMAIN 393 397 POLY-PRO.
FT DOMAIN 571 590 POLY-ASN.
FT DOMAIN 1808 1811 POLY-THR.
FT DOMAIN 1891 1897 POLY-GLN.
FT CONFLICT 2275 2275 R -> W (IN REF. 2).
FT SEQUENCE 2278 AA: 257417 MW: 1A0A30E13165DE41 CRC64;

Query Match 4.78; Score 112.5; DB 1; Length 2278;
Best Local Similarity 19.08; Pred. No. 1.9;
Matches 88; Conservative 57; Mismatches 163; Indels 155; Gaps 18;

OY KRRKRKSYGVYVNY-----KDC-----NNTLQKTEVRECCSPFGSMLCGSRGL 103
DB 221.RKDKDTNNNGVLKSEYWMKDESSKEFCSCGTFMTRKHKRCICGQIFCSCPLLIDGD 280
OY 104.QFHLNNSHDLFEFFFKLFEEYQYVNVSVKLSNLFEEBGSDDDKFEFDSLCSPRRKR-RQ 162
DB 281.RF---GCHAKRVCYNYCEIADYI-----EDSSDEENDSTKQL-NEPKRSRS 324
OY 163.RGGRNNTRRRLKVCFLPLDPSLTLNGT-----ENGITLLDNGKRLGYDEAT 208
DB 325.RSNTINFYSHSHSLHLISQDNNHGTDLDPVATDMPQQONEVYLLDDDD-----VQSI 379
OY 209.ELAGQ-----FEMTNSIDPAILAHSSLDGACAVILTSEAVPATK---TRKLSAEKSEARSH 261

```

Db	380	MTSGEDSKLEISIPPPPPMAA-----IPATKQGSLEISFSDENDRA- 422
Qy	262	LLIQRQFYSHERVOPMALEQVMSDRDSEYDDVDADFEORQIMLDFVDVNRKDEQFMH 321
Db	422	LHYQDDNPGRHHHLDVPTPTRTIRDDNISHYDTN-----NSTSLRPHYNNNSTITINN 476
Qy	322	LMS-----FPRQRVIADGIISMACAFSRYEKE 352
Db	477	LNNTTSSNNSNYNTNNSNSINNPNSHLRSRISPHYVSSSVKNSDSSNSATPASS----- 530
Qy	353	LHRYSSLEWCWRLFLIKLWNHGLVDSATINNCITLENC--RN-----SSD 396
Db	531	-----AQSSILDPNRLITIGYAHARNYKFKRYNSKQPSQOND 566
RESULT 13		
SWI1_YEAST		
SWI1_YEAST	STANDARD;	PRT; 1314 AA.
AC	P09547;	
DT	01-MAR-1989 (Rel. 10, Last sequence update)	
DT	01-MAR-1989 (Rel. 10, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Transcription regulatory protein SWI1 (SWI/SNF complex component SWI1)	
DE	(Transcription regulatory protein AdR6) (Regulatory protein GAN3).	
GN	AD66 OR SWI1 OR GAN3 OR YPL016W OR LPL1.	
OS	Saccharomyces cerevisiae (Baker's yeast).	
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.	
OX	NCBI_TaxID=4932;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RP	MEDLINE=89057455; PubMed=3143101;	
RA	O'Hara P.J., Horowitz H., Eichinger H., Young E.T.;	
RT	"The yeast AdR6 gene encodes homopolymeric amino acid sequences and a	
RL	potential metal-binding domain."	
RL	Nucleic Acids Res. 16:10153-10170(1988).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=S288c / AB972;	
RX	MEDLINE=97313271; PubMed=9169875;	
RA	Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgorge W.,	
RA	Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,	
RA	Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,	
RA	Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,	
RA	Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,	
RA	Duncan M., Floeth M., Fortin N., Friesen J.D., Filtz C., Goffeau A.,	
RA	Hall J., Heblung U., Heumann K., Hilbert H., Hillier L.,	
RA	Hunkle-Smith S., Hyman R., Johnson M., Kallman S., Kleene K.,	
RA	Komp C., Kurdi O., Laskari D., Lew H., Lin A., Lin D., Louis E.J.,	
RA	Marathe R., Messenguy F., Mewes H.-W., Mittlepat S., Moestl D.,	
RA	Mueller-Auer S., Namath A., Nentwich U., Oetner P., Pearson M.,	
RA	Petel F.X., Pohl T.M., Purnelle D., Schater M., Schafie M.,	
RA	Ureterescu B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,	
RA	Vierstraete L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,	
RA	Walsh S.V., Wambut R., Wang Y., Medler E., Wedler H., Winnett E.,	
RA	Zhong W.W., Zollner A., Vo D.H., Hani J.;	
RL	"The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."	
RL	Nature 387:103-105(1997).	
RN	[3]	
RP	CHARACTERIZATION.	
RX	MEDLINE=92154671; PubMed=1339306;	
RA	Peterson C.L., Herskowitz I.;	
RT	"Characterization of the yeast SWI1, SWI2, and SWI3 genes, which	
RT	encode a global activator of transcription."	
RL	Cell 68:573-583(1992).	
CC	-I- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE SWI/SNF	
CC	COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER	
CC	OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE	
CC	BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.	


```
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- DEVELOPMENTAL STAGE: AFTER FRUITING BODIES HAVE BEEN FORMED AND
CC DURING GERMINATION.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 CHASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M87278; AAA33164.1; -.
DR PDB: 1AMW; 28-JAN-98.
DR DictyDB; DD04007; acgA.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS50839; CHASE; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Lyase; CAMP synthesis; Transmembrane; Germination; signal-anchor;
KW 3D-structure.
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 42 858 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 86 317 CHASE.
FT DOMAIN 396 526 GUANYLATE CYCLASE.
FT DOMAIN 738 848 ASN-RICH.
FT DOMAIN 835 848 POLY-ASN.
SQ SEQUENCE 858 AA; 98413 MW; B483FB255289E65D CRC64;

Query Match 4.7%; Score 111.5; DB 1; Length 858;
Best Local Similarity 19.2%; Pred. No. 0.64;
Matches 76; Conservative 57; Mismatches 135; Indels 127; Gaps 15;

QY 104 QPHLNS-HDLFEPEKLFEEYQIVNVSVK--LNSLFEEEGSDDDKFEPPSLCSKPRK 159
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 533 QIHVSDRYQLGKEDFNFSERCDLIHVKGGMKTWYLMGKKSDPSLKKDFS-----RS 587
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 160 RRORGRRNTRRLKVCFLPLDSPSLTNGTENGITILLDGNRGLGYPEATELAGOFEMTSN 219
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 588 RVO-----PSLFNRKQSHVCI-----YPEPPSGIQLALNIENN 620
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 220 IPPAIHSSLDAG---AKVILTSEAVVPATKTRKLSAERSPARSHLLLOKROFYHSHRV 275
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 621 L-----NNTDAGCENCCKILKTYAYSPDST-----SNYYTHGDDN 657
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 276 QPMALQVMSDR-DSEDEVDVADFEEDROMLIDFVYNKDEKQPMH--LMSFVYRKQRY 332
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 658 SPPEPSLNSNDLIDSEFYHDDPFPSDSNVGYHDTSKDIKEDENQNETLLFNOQLKKQ 717
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 333 IADGHSWACEAFSRYEKELEHRYSLFWCWRFLFLIKLMNHGLVDSATLNNCNTILLENCR 392
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 718 IEN-----IQRDLSLND-----IEAKILNNNNNNNNNNNNNNNNNTNPNNN 758
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 393 -----NSSDTT---NNNNSVDRPSDSNTNNNNIVD----- 422
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 759 NNDIYNNSDNNYNNNNNNNSDKLENNDDGNNNNINDNNYKSTNENNIKSKTLFQDSKL 818
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 423 -----PNDINKNNVNDKNNNSRDKVYK 445
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 819 INDIKAKENCDDNDDNNNNNNNNNNNNNNNNNNNNNNDENVE 853
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Search completed: June 13, 2003, 15:36:11
Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 17:45:40 ; Search time 4487 Seconds

(without alignments)
11168.939 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctctcaattgtct.....aatgattctgtcttaact 1722

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1722	100.0	1722	6	AX032817	AX032817 Sequence
2	1722	100.0	1722	6	AX032888	AX032888 Sequence
3	1720.4	99.9	1722	6	AX032934	AX032934 Sequence
4	1678.6	97.5	1700	8	AF284500	AF284500 Arabidops
5	1646.6	95.6	1722	8	AF284501	AF284501 Arabidops
6	1569.6	91.1	1715	6	AX032818	AX032818 Sequence
7	1569.6	91.1	1715	6	AX032891	AX032891 Sequence
8	1568.6	91.1	1721	8	AY034902	AY034902 Arabidops
9	1537.6	89.3	1354	6	AX032894	AX032894 Sequence
10	1277.4	74.2	1354	6	AY063047	AY063047 Arabidops
11	412.2	23.9	6338	6	AX032890	AX032890 Sequence
12	407.8	23.7	91660	8	ATRP51E2	AF180942 Arabidops
13	319.4	18.5	5895	6	AX032893	AX032893 Sequence
14	319.4	18.5	199548	8	ATCHRIV45	AL161545 Arabidops
15	319.4	18.5	201471	8	ATPCA7	Z97342 Arabidops
16	291.4	16.9	2279	8	AB053171	AB053171 Arabidops
17	236.2	13.7	2384	8	AF096096	AF096096 Arabidops
18	230.8	13.4	558	6	AX032900	AX032900 Sequence
19	163.8	9.5	325	6	AX032898	AX032898 Sequence
20	126.6	7.4	463	6	AX032896	AX032896 Sequence
21	78.4	4.6	35276	2	AC115611	AC115611 Dictyoste
22	75.2	4.4	256774	2	AC116964	AC116964 Dictyoste
23	73.4	4.3	106434	3	AC117080	AC117080 Dictyoste
24	73.4	4.3	147192	2	AC116925	AC116925 Dictyoste
25	73.2	4.3	79554	8	NCB11823	AL669991 Neurospor
26	72.8	4.2	141008	2	AC126897	AC126897 Rattus no
27	72.8	4.2	144022	2	AC109949	AC109949 Rattus no
28	72.2	4.2	185424	2	AC096166	AC096166 Rattus no
29	72.2	4.2	166918	2	AC128567	AC128567 Rattus no
30	72.2	4.2	237563	2	AC102954	AC102954 Rattus no
31	71.8	4.2	7921	3	AF153362	AF153362 Dictyoste
32	71.8	4.2	97683	2	AC116348	AC116348 Dictyoste
33	71.8	4.2	212383	2	AC119030	AC119030 Rattus no
34	71.6	4.2	57906	8	NCB7J19	AL389899 Neurospor
35	71.6	4.1	656	3	AF453312	AF453312 Cotesia c
36	71.4	4.1	8045	6	AX034425	AX034425 Sequence
37	71.4	4.1	156533	2	AC117070	AC117070 Dictyoste
38	71.4	4.1	229502	2	AC126871	AC126871 Rattus no
39	70.6	4.1	153359	2	AC096792	AC096792 Rattus no
40	70.6	4.1	165337	2	AC096212	AC096212 Rattus no
41	70.6	4.1	176351	2	AC117007	AC117007 Rattus no
42	70.6	4.1	179380	2	AC120325	AC120325 Rattus no
43	70.4	4.1	2790	3	DD301670	AJ301670 Dictyoste
44	70.4	4.1	6016	8	AY077730	AY077730 Pedospora
45	70.4	4.1	53440	2	AC117078	AC117078 Dictyoste

ALIGNMENTS

RESULT 1	AX032817	1722 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	AX032817				
DEFINITION	Sequence 57 from Patent WO0046558.				
ACCESSION	AX032817				
VERSION	AX032817.1	GI:10279793			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Dean,C., West,J. and Johanson,U.				
TITLE	Plant gene				
JOURNAL	Patent: WO 0046358-A 57 10-AUG-2000;				
	DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;				

Pred. No. is the number of results predicted by chance to have a

JOHANSON URBAN (SE)
Location/Qualifiers
1. 1722

/organism="synthetic construct"
/db.xref="taxon:32630"
/note="Landsberg erecta VRN2 cDNA"

BASE COUNT 538 a 339 c 348 g 497 t
ORIGIN

Query Match 100.0%; Score 1722; DB 6; Length 1722;
Best Local Similarity 100.0%; P-Val. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCTCTTCAATTTCTGCTCTCTCTTACAGACCAATCGGTGTTTCGACCTTT 60
DB 1 CAAGCTCTTCAATTTCTGCTCTCTCTTACAGACCAATCGGTGTTTCGACCTTT 60
QY 61 CAGGCTCAATCCAAAGACATCTATATAGCATATTCAGAGAGGCGGTTCTAATGTT 120
DB 61 CAGGCTCAATCCAAAGACATCTATATAGCATATTCAGAGAGGCGGTTCTAATGTT 120
QY 121 GCATTGAGTTTATCCGCTATGAGAGGAAATTTCTAATTTAGGGAGCGCTCAGAGTTTG 180
DB 121 GCATTGAGTTTATCCGCTATGAGAGGAAATTTCTAATTTAGGGAGCGCTCAGAGTTTG 180
QY 181 CAGTACTCATTAATCGGCTCTGACGTTGTGAGTGAATTAACAAAGAAATGTGAGGC 240
DB 181 CAGTACTCATTAATCGGCTCTGACGTTGTGAGTGAATTAACAAAGAAATGTGAGGC 240
QY 241 AAGATGTGCGGCGAAATCTCACCGAGAGAGATGATTTCACTGATGAGATCTCTTGA 300
DB 241 AAGATGTGCGGCGAAATCTCACCGAGAGAGATGATTTCACTGATGAGATCTCTTGA 300
QY 301 TATATTTGAACCTGTGACATATTAACATCTTTACCTTCGCTCTAGGCAACCAT 360
DB 301 TATATTTGAACCTGTGACATATTAACATCTTTACCTTCGCTCTAGGCAACCAT 360
QY 361 CGTTTCTTCCAGATGCTTGAACATAAAATTTGAGCAAAAGCCAAAGAAAGTCAAGAT 420
DB 361 CGTTTCTTCCAGATGCTTGAACATAAAATTTGAGCAAAAGCCAAAGAAAGTCAAGAT 420
QY 421 CTACTGAGATGTTTCAACTATAAGATGTATTAACATCTTACAGAAACGTAAG 480
DB 421 CTACTGAGATGTTTCAACTATAAGATGTATTAACATCTTACAGAAACGTAAG 480
QY 481 TTAGGAGAGATGTTCTGCTGCTATTTGCTATGATGTGAGTTCAGAGGCTGC 540
DB 481 TTAGGAGAGATGTTCTGCTGCTATTTGCTATGATGTGAGTTCAGAGGCTGC 540
QY 541 AATTCATTTGAATTCATCTCATGATTAATTTGAATTTGAGTTCAGACCTTTGAGAAAT 600
DB 541 AATTCATTTGAATTCATCTCATGATTAATTTGAATTTGAGTTCAGACCTTTGAGAAAT 600
QY 601 ACCAGACAGTAAATGTTCTGTAAAACTTAATTCCTCATATTTGAGAGAAAGAGTG 660
DB 601 ACCAGACAGTAAATGTTCTGTGTAAAACTTAATTCCTCATATTTGAGAGAAAGAGTG 660
QY 661 ATGACGATTAATTTAGGCTTCTCTCTCTGCTGAGAACTCGTAAGCGAGCAAGAG 720
DB 661 ATGACGATTAATTTAGGCTTCTCTCTCTGCTGAGAACTCGTAAGCGAGCAAGAG 720
QY 721 GTGGGAGAAATTAACCGAGAGACTTAAGATGCTTTTACCGTTGAGATTCACCGAGTT 780
DB 721 GTGGGAGAAATTAACCGAGAGACTTAAGATGCTTTTACCGTTGAGATTCACCGAGTT 780
QY 781 TAACTAATGAGCAGAGAAATGAAATCAACCTTAATTAATGAGAAACCGTGTAGAT 840
DB 781 TAACTAATGAGCAGAGAAATGAAATCAACCTTAATTAATGAGAAACCGTGTAGAT 840
QY 841 ATCCGAGGCAACAGAGCTGCTGAGCAATTTGAGATGAGCAGCAACATCTCCACGACA 900
DB 841 ATCCGAGGCAACAGAGCTGCTGAGCAATTTGAGATGAGCAGCAACATCTCCACGACA 900

QY 901 TAGCCACCTTCTCTGACGCTGCTGCTAAAGTTATTTATGACAGAGAGAGTGTGCTC 960
DB 901 TAGCCACCTTCTCTGACGCTGCTGCTAAAGTTATTTATGACAGAGAGAGTGTGCTC 960
QY 961 CTGCTCTCTAAGCAAGAAAGTATCTGCTGAGGATGAGAGGCTAGAGACCTTACTTC 1020
DB 961 CTGCTCTCTAAGCAAGAAAGTATCTGCTGAGGATGAGAGGCTAGAGACCTTACTTC 1020
QY 1021 TTGAGAAAGCCCAATTTCTATCTTCTCAGACATCTCCAGCAATGCGCTTGAGCAATGA 1080
DB 1021 TTGAGAAAGCCCAATTTCTATCTTCTCAGACATCTCCAGCAATGCGCTTGAGCAATGA 1080
QY 1081 TGTCTGACCGGGATGACGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1140
DB 1081 TGTCTGACCGGGATGACGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1140
QY 1141 AGATGCTGTATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 AGATGCTGTATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ACTGCTTTGAGAAAGCAAGAGGTTATAGAGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 ACTGCTTTGAGAAAGCAAGAGGTTATAGAGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CATTTCAAGATTTTACAGAGAAAGAGTGCACGCTTCTCATGATGATGATGATGATGATGAT 1320
DB 1261 CATTTCAAGATTTTACAGAGAAAGAGTGCACGCTTCTCATGATGATGATGATGATGATGAT 1320
QY 1321 GATTTCTTTGATTAACATATGAGAACATGAGATCTTGCACCTGATGATGATGATGATGAT 1380
DB 1321 GATTTCTTTGATTAACATATGAGAACATGAGATCTTGCACCTGATGATGATGATGATGAT 1380
QY 1381 GCATATCAATCTCGAGAAATGCGGTAATGCTCAGACACACACACACACACACACACACACA 1440
DB 1381 GCATATCAATCTCGAGAAATGCGGTAATGCTCAGACACACACACACACACACACACACACA 1440
QY 1441 ACAGTGTGATGCTGCCAGTCACTCAACACACACACACACACACACACACACACACACACACA 1500
DB 1441 ACAGTGTGATGCTGCCAGTCACTCAACACACACACACACACACACACACACACACACACACA 1500
QY 1501 ATGACATTAACCAAGAAAGATGTTGCAACAGAGACATTAACAGAGAGAGAGAGAGAGAGAG 1560
DB 1501 ATGACATTAACCAAGAAAGATGTTGCAACAGAGACATTAACAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 TTAAGAGAGAAATGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1561 TTAAGAGAGAAATGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 TTTCTTAAGAAATGTTTGAAGCAACAAATTTTAAATGATGATGATGATGATGATGATGAT 1680
DB 1621 TTTCTTAAGAAATGTTTGAAGCAACAAATTTTAAATGATGATGATGATGATGATGATGAT 1680
QY 1681 TACATTTTATGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1722
DB 1681 TACATTTTATGTTAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1722

RESULT 2
AX032888
LOCUS 1722 bp DNA linear PAT 22-SEP-2000
DEFINITION Sequence 1 from Patent WO0044918.
ACCESSION AX032888
VERSION AX032888.1 GI:10279825
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1722)
REFERENCE
Dean, C. and Gendall, A.
Methods and means for modification of plant characteristics using
the vernalization gene vrn2

the vernalization gene *Vrn2*
 Patent: WO 0044918-A 47 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)

FEATURES
 source Location/Qualifiers
 1..1722
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"

BASE COUNT 539 a 339 c 347 g 497 t

ORIGIN

Query Match 99.98%; Score 1720.4; DB 6; Length 1722;
 Best Local Similarity 99.98%; Pred. No. 0;
 Matches 1721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGCTTCTCAATTTGCTTCTCTCTTACACAGCCATCGTGTTCGAGCTT 60
 Db 1 CAAGCTTCTCAATTTGCTTCTCTCTTACACAGCCATCGTGTTCGAGCTT 60
 QY 61 CAGGCTCAATCCAGACATCTATATAGCATATGAGAGAGGGGTTCTATATGT 120
 Db 61 CAGGCTCAATCCAGACATCTATATAGCATATGAGAGAGGGGTTCTATATGT 120
 QY 121 GGATGAGTTTATGCTATGAGTGAAGAAATTTCAATTTAGGGAGGCTCAGAGTTG 180
 Db 121 GGATGAGTTTATGCTATGAGTGAAGAAATTTCAATTTAGGGAGGCTCAGAGTTG 180
 QY 181 CACTAATCTCAATATCGCTCTTGACGTTGTGAGTGAATTTGAACAAGAAATGTAGGC 240
 Db 181 CACTAATCTCAATATCGCTCTTGACGTTGTGAGTGAATTTGAACAAGAAATGTAGGC 240
 QY 241 AGAATTTGCGCGAAATCTCCACCGAGAGAAAGTATTTCAACGATGAGAAATCTTTGA 300
 Db 241 AGAATTTGCGCGAAATCTCCACCGAGAGAAAGTATTTCAACGATGAGAAATCTTTGA 300
 QY 301 TATATGTAAACCTGTGAGATATATTAACATCTTTCACTCGCTCTGAGCAACCAT 360
 Db 301 TATATGTAAACCTGTGAGATATATTAACATCTTTCACTCGCTCTGAGCAACCAT 360
 QY 361 CGTTTCTCCAGATGCTGAACCTGAACAAATTTGAGGCAAGGCAAGAAAGTCAAGAT 420
 Db 361 CGTTTCTCCAGATGCTGAACCTGAACAAATTTGAGGCAAGGCAAGAAAGTCAAGAT 420
 QY 421 CTACTGGAGTGTAGTTTCAACATTAAGGATGTAAATACACATTTACAGAAATCTGAG 480
 Db 421 CTACTGGAGTGTAGTTTCAACATTAAGGATGTAAATACACATTTACAGAAATCTGAG 480
 QY 481 TTAGGAGGATTTGTTCTGTCCATTTTCTATGCTATGCTATGCTTCAAGGGCTGC 540
 Db 481 TTAGGAGGATTTGTTCTGTCCATTTTCTATGCTATGCTATGCTTCAAGGGCTGC 540
 QY 541 AATTCATTTGAATTCATCTCATGATTAATTTGAATTTGAATTTGAATTTGAATTTGAAT 600
 Db 541 AATTCATTTGAATTCATCTCATGATTAATTTGAATTTGAATTTGAATTTGAATTTGAAT 600
 QY 601 ACCAGACAGTTAATTTCTCTGAACCTAATTTCTCTCATTTGAGGAGAAAGAGAGTG 660
 Db 601 ACCAGACAGTTAATTTCTCTGAACCTAATTTCTCTCATTTGAGGAGAAAGAGAGTG 660
 QY 661 ATGAGGATTAATTTGAGGCTTCTCTCTGCTGAAACCTGTGAAGGAGACAAAGAG 720
 Db 661 ATGAGGATTAATTTGAGGCTTCTCTCTGCTGAAACCTGTGAAGGAGACAAAGAG 720
 QY 721 GTGGCAGAAATTAACACAGAGACTTAAGATAGCTTTTACCGTTGATTCACCCAGTT 780
 Db 721 GTGGCAGAAATTAACACAGAGACTTAAGATAGCTTTTACCGTTGATTCACCCAGTT 780
 QY 781 TAACATATGACACAGAAATGGAATCACCCTACTTAATGATGGAACCGGTGTAGAT 840
 Db 781 TAACATATGACACAGAAATGGAATCACCCTACTTAATGATGGAACCGGTGTAGAT 840
 QY 841 ATCCGAGGCAACAGAGTGTGTGCAAAATTTGAGTACACAGAAATTTCCACCA 900
 Db 841 ATCCGAGGCAACAGAGTGTGTGCAAAATTTGAGTACACAGAAATTTCCACCA 900

Db 841 ATCCGAGGCAACAGAGTGTGTGCAAAATTTGAGTACACAGAAATTTCCACCA 900
 QY 901 TAGCCCACTCTTCCTGAGAGTGTGTGCAAAATTTGAGTACACAGAAATTTCCACCA 960
 Db 901 TAGCCCACTCTTCCTGAGAGTGTGTGCAAAATTTGAGTACACAGAAATTTCCACCA 960
 QY 961 CTGCTACTAGACAGAAAGTATCTGCTGAGGATCAGAGGCTAGAACCCACTTACTTC 1020
 Db 961 CTGCTACTAGACAGAAAGTATCTGCTGAGGATCAGAGGCTAGAACCCACTTACTTC 1020
 QY 1021 TTCAGAAAGCCCAATTTCTATCATCTCAGAGTTCACAGCAATGAGGCTTGAGCAATGA 1080
 Db 1021 TTCAGAAAGCCCAATTTCTATCATCTCAGAGTTCACAGCAATGAGGCTTGAGCAATGA 1080
 QY 1081 TGTCTGACCGGAGATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 Db 1081 TGTCTGACCGGAGATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1141 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Db 1141 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 ACTGCTTTGTAAGAAACAAAGGTTTATACAGATGTCATATCTCTGGCATGTGAG 1260
 Db 1201 ACTGCTTTGTAAGAAACAAAGGTTTATACAGATGTCATATCTCTGGCATGTGAG 1260
 QY 1261 CATTTTCAAGATTTTACAGAAAGTTCAGACCGTTTACTCATCATCTCTCTGAGTTGA 1320
 Db 1261 CATTTTCAAGATTTTACAGAAAGTTCAGACCGTTTACTCATCATCTCTCTGAGTTGA 1320
 QY 1321 GATTTGTTTATTTAACTATGAAACCATGACTTGTGCACTGACGACCATCAACACT 1380
 Db 1321 GATTTGTTTATTTAACTATGAAACCATGACTTGTGCACTGACGACCATCAACACT 1380
 QY 1381 GCAATATCATCTCTGAGAAATTCGCTATATGCTCAAGACACACCAACCAACAAACA 1440
 Db 1381 GCAATATCATCTCTGAGAAATTCGCTATATGCTCAAGACACACCAACCAACAAACA 1440
 QY 1441 ACAGTGTGATGCTGCTGAGAAATTCGCTATATGCTCAAGACACACCAACCAACAAACA 1500
 Db 1441 ACAGTGTGATGCTGCTGAGAAATTCGCTATATGCTCAAGACACACCAACCAACAAACA 1500
 QY 1501 ATGACATTAACACAGAAAGTATGTAACAAAGGACATTAACACAGAAAGTAA 1560
 Db 1501 ATGACATTAACACAGAAAGTATGTAACAAAGGACATTAACACAGAAAGTAA 1560
 QY 1561 TTAATTTAGAAATCTCGGCTTTTATGATACGATTTATCGATTTGTAATTTATCTTC 1620
 Db 1561 TTAATTTAGAAATCTCGGCTTTTATGATACGATTTATCGATTTGTAATTTATCTTC 1620
 QY 1621 TTTCTTAAATAATTTAGGAGCAACAATTTTATATGTTAGTATCAACTGAT 1680
 Db 1621 TTTCTTAAATAATTTAGGAGCAACAATTTTATATGTTAGTATCAACTGAT 1680
 QY 1681 TACATTTTATGTTAAAAAATGATGATGATGATGATGATGATGATGATGATGATG 1722
 Db 1681 TACATTTTATGTTAAAAAATGATGATGATGATGATGATGATGATGATGATGATG 1722

RESULT 4
 AF284500 1700 bp mRNA linear PLN 21-NOV-2001
 LOCUS AF284500
 DEFINITION Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
 cds.
 ACCESSION AF284500
 VERSION AF284500.1 GI:16945787
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 1700)

AUTHORS Gendall, A.R., Levy, Y.Y., Wilson, A. and Dean, C.
TITLE The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis
JOURNAL Cell 107 (4), 525-535 (2001)
MEDLINE 21575875
PUBMED 11719192
REFERENCE 2 (bases 1 to 1700)
AUTHORS Gendall, A.R., Levy, Y.Y. and Dean, C.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK

FEATURES
source
 1..1700
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="IV"
 /map="adjacent to RPP5"
 /note="ecotype: Landsberg erecta"
 1..1700
 /gene="VRN2"
 227..1564
 /gene="VRN2"
 /codon_start=1
 /product="vernalization 2 protein"
 /protein_id="AAL32135.1"
 /db_xref="GI:16945788"
 /translation="MCRONRAKSSPEVYSIDENILTYCKPRLNTPHLSIGNS
 FLPRCLNTKIKAKRRKRSSTGVYFNKYKDCNNTLOKTEVEDDCSCPPSMLGSGK
 LQFHLNNSHDLFEFEKLESEYQTVNVSXKLSFLEEKSDDDDFEPLSLSKPKR
 RQGRNNTNRLKVCFLPLDPSLITNGENGLTLNDGNRGLGYEPALAGOFKETS
 NIPPAIASHSLDAGAKVILTSADVVPATKRLSERSRSHLLLOKQFYHSHRVQ
 PMALQVMSDEDEVDVADFEDRQMLDQFVDVNDKDEKQFMHLMNSFVKORVIA
 DSHISMACAFSRYFEKELHRYSLFMCWRFLIKLMNHGLVDSATINNCNTILNCR
 NSDDTTTNNNSVDRPDSNTNNNNIYDHPNDINKNNVDKDNNSRDKYIK"
 853
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="a"
 1195
variation
 /gene="VRN2"
 /note="vrn2-1 mutation results in TGG Trp to TGA stop"
 /replace="a"
 1399
variation
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="g"
 1487
variation
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="a"
 1488
variation
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="g"
 BASE COUNT 533 a 334 c 347 g 486 t
 ORIGIN
 Query Match 97.5%; Score 1678.6; DB 8; Length 1700;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1681; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

140 GACGTAGGGAATTCCTAATTTAGGGAGGCGCTCAGATTTGCATCTAATTCATTAATCGGC 199
 |||||||
 136 GACGTAGGGAATTCCTAATTTAGGGAGGCGCTCAGATTTGCATCTAATTCATTAATCGGC 195
 |||||||
 200 TCTTGAAGCTGTGTGAGTGTAAATGGAACAAGAAATGTGTAGCAGAAATTTGGCGCGGAATC 259
 |||||||
 196 TCTTGAAGCTGTGTGAGTGTAAATGGAACAAGAAATGTGTAGCAGAAATTTGGCGCGGAATC 255
 |||||||
 260 CTCACCGGAGGAAGATTTTCAACTGATGAGAAATCTCTTGAATATATTTGAACCTGTGG 319
 |||||||
 256 CTCACCGGAGGAAGATTTTCAACTGATGAGAAATCTCTTGAATATATTTGAACCTGTGG 315
 |||||||
 320 ACTATATTAACATCTTCCACCTCGCTCTCTAGCAACCCATCGTCTTCTTCCAAAGATGCTT 379
 |||||||
 316 ACTATATTAACATCTTCCACCTCGCTCTCTAGCAACCCATCGTCTTCTTCCAAAGATGCTT 375
 |||||||
 380 GAACATCAAAATTTGAGCAAAAGCGCAAAAGAAAGTCAAGATCTACTGGAATGGTACTTTT 439
 |||||||
 376 GAACATCAAAATTTGAGCAAAAGCGCAAAAGAAAGTCAAGATCTACTGGAATGGTACTTTT 435
 |||||||
 440 CAACATTAAGGATTTGTAATTAACACATTTACAGAAACTGAAGTTAGGAGAGATTTCTTGG 499
 |||||||
 436 CAACATTAAGGATTTGTAATTAACACATTTACAGAAACTGAAGTTAGGAGAGATTTCTTGG 495
 |||||||
 500 TCCATTTTCTCTATGCTATGTGTGTGAGCTTCAAGGGGCTGCAATTTCAATTTGAATTCATC 559
 |||||||
 496 TCCATTTTCTCTATGCTATGTGTGTGAGCTTCAAGGGGCTGCAATTTCAATTTGAATTCATC 555
 |||||||
 560 TCATGATTTATTTGAATTTGAGTCAAGCTTTTGAAGAATACAGACAGTTAAATGTTTC 619
 |||||||
 556 TCATGATTTATTTGAATTTGAGTCAAGCTTTTGAAGAATACAGACAGTTAAATGTTTC 615
 |||||||
 620 TGTAAACCTAATTCCTTCAATATTTAGGAAGAAGAAAGTGAATGACGATTAATTTGAGGC 679
 |||||||
 616 TGTAAACCTAATTCCTTCAATATTTAGGAAGAAGAAAGTGAATGACGATTAATTTGAGGC 675
 |||||||
 680 CTTCCTCTCTCTCTCGAAGACCTGTAAAGGAGAGACAAAGAGGTGGCAGAAATTAACACAG 739
 |||||||
 676 CTTCCTCTCTCTCTCGAAGACCTGTAAAGGAGAGACAAAGAGGTGGCAGAAATTAACACAG 735
 |||||||
 740 GAGACTTAAAGTATGCTTTTACCGTTGATTCACCAAGTTTAACATTAATGCGACAGAA 799
 |||||||
 736 GAGACTTAAAGTATGCTTTTACCGTTGATTCACCAAGTTTAACATTAATGCGACAGAA 795
 |||||||
 800 TGGAAATCACCTACTTAATGATGGAAGCAAGCTGTGTAGATATCCCGAGGCAACAGAGCT 859
 |||||||
 796 TGGAAATCACCTACTTAATGATGGAAGCAAGCTGTGTAGATATCCCGAGGCAACAGAGCT 855
 |||||||
 860 TGTGTGACATTTGAGATGACCAAGCAATTCACCAAGCCATAGGCCACTCTTCTCTGGA 919
 |||||||
 856 TGTGTGACATTTGAGATGACCAAGCAATTCACCAAGCCATAGGCCACTCTTCTCTGGA 915
 |||||||
 920 CGCTGTGCTAAAGTTATATTTGACAAAGCAAGCTGTGTCTCTCTCTCTCTCTCTCTCTCT 979
 |||||||
 916 CGCTGTGCTAAAGTTATATTTGACAAAGCAAGCTGTGTCTCTCTCTCTCTCTCTCTCTCT 975
 |||||||
 980 GTTATCTGTGAGCGCATGAGAGCTTAGAGCCACCTACTCTTCTGAAACGCCAATTTCTA 1039
 |||||||
 976 GTTATCTGTGAGCGCATGAGAGCTTAGAGCCACCTACTCTTCTGAAACGCCAATTTCTA 1035
 |||||||
 1040 TCATTTCTCAGACAGTCCACCAATAGCGCTTGTAGCAAGTAATGTCTGACCGGATAGCGA 1099
 |||||||
 1036 TCATTTCTCAGACAGTCCACCAATAGCGCTTGTAGCAAGTAATGTCTGACCGGATAGCGA 1095
 |||||||
 1100 GGATGAAGTCATGACGATGTTCAGATTTTGAAGATGCCAGATGCTTGAATGACTTGT 1159
 |||||||
 1096 GGATGAAGTCATGACGATGTTCAGATTTTGAAGATGCCAGATGCTTGAATGACTTGT 1155
 |||||||
 1160 GGATGAAGTCATGACGATGTTCAGATTTTGAAGATGCCAGATGCTTGAATGACTTGT 1219
 |||||||
 1156 GGATGAAGTCATGACGATGTTCAGATTTTGAAGATGCCAGATGCTTGAATGACTTGT 1215
 |||||||
 1220 AAGGTTATAGCAGATGTCTATCTCTTGGGATGTGAAGCATTTTCAAGATTTTACGA 1279

Db	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600	1601	1602	1603	1604	1605	1606	1607	1608	1609	1610	1611	1612	1613	1614	1615	1616	1617	1618	1619	1620	1621	1622	1623</
----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

BASE COUNT	540 a	338 c	353 g	491 t	
ORIGIN					
Query Match					
Best Local Similarity	95.64;	Score 1646.6;	DB 8;	Length 1722;	
Matches 1681;	Conservative	0;	Mismatches	4;	Indels 22;
					Gaps 1;
Db	20	TTGCTCTCTTACACAGCCAAATCGTGTGTTTGGAGGTTTCAGAGCTTCATCCAGACA	79		
Db	16	TTCTCTCTCTTACACAGCCAAATCGTGTGTTTGGAGGTTTCAGAGCTTCATCCAGACA	79		
QY	80	TTCTATATTAAGCATATTCAGAAAGGGCGGTCTTAATTTGTCATTTAGAGTTTATCGCTAT	75		
Db	76	TTCTATATTAAGCATATTCAGAAAGGGCGGTCTTAATTTGTCATTTAGAGTTTATCGCTAT	139		
QY	140	GACGTAGGAAATTTCTAATTTAGGAGGGCTCTAGAGTTTGCACTAACTCTAATATCGGC	135		
Db	136	GACGTAGGAAATTTCTAATTTAGGAGGGCTCTAGAGTTTGCACTAACTCTAATATCGGC	199		
QY	200	TTCTGACCTTGTAGGTATTTGAACAAGAATGTATAGCAGAATTTGCGCGCAAAATC	195		
Db	196	TTCTGACCTTGTAGGTATTTGAACAAGAATGTATAGCAGAATTTGCGCGCAAAATC	259		
QY	260	CTACCGGAGAGATGTTTCACTCATATGAGAAATCTTGATATATTTGAAACCTGTATCG	255		
Db	256	CTACCGGAGAGATGTTTCACTCATATGAGAAATCTTGATATATTTGAAACCTGTATCG	319		
QY	320	ACTATATTAACATCTTTCACCTTCGCTCTGAGCAACCAATCGTTTCTCCAAATCGTT	315		
Db	316	ACTATATTAACATCTTTCACCTTCGCTCTGAGCAACCAATCGTTTCTCCAAATCGTT	379		
QY	380	GAACATCAAAATTTGGAGCAAGCGGAAAGAAAGTCAAGATCTACTGGATGGTAGTTT	375		
Db	376	GAACATCAAAATTTGGAGCAAGCGGAAAGAAAGTCAAGATCTACTGGATGGTAGTTT	439		
QY	440	CAACTATTAAGATTTGTAATTAACACATTAACAGAAATCTGAAATTAGGAGAGATGTCTTG	435		
Db	436	CAACTATTAAGATTTGTAATTAACACATTAACAGAAATCTGAAATTAGGAGAGATGTCTTG	499		
QY	500	TCGATTTTGGCTATAGCATGTGGTAGTCTCA-----AGGGGC	495		
Db	496	TCGATTTTGGCTATAGCATGTGGTAGTCTCA-----AGGGGC	553		
QY	538	TGCATTTTCAATTTGAATTCATCTCATGATTAATTTGAATTTAGTTCAAGCTTTTCGAG	555		
Db	536	TGCATTTTCAATTTGAATTCATCTCATGATTAATTTGAATTTAGTTCAAGCTTTTCGAG	597		
QY	598	AATACAGACAGTAAATGTTCTGTAATAACTAATTCCTTCATATTTGAGGAAGAAAGAA	615		
Db	616	AATACAGACAGTAAATGTTCTGTAATAACTAATTCCTTCATATTTGAGGAAGAAAGAA	657		
QY	658	GTGATGACATTAATTTTGGAGCGCTCTCTCTCTCTGAAACCTGTAACGGAGACAAA	675		
Db	676	GTGATGACATTAATTTTGGAGCGCTCTCTCTCTCTGAAACCTGTAACGGAGACAAA	717		
QY	718	GAGGTGGAGAAATTAACACAGGAGACTTAAAGTATGCTTTTACGGTTGAGATTCACCA	735		
Db	736	GAGGTGGAGAAATTAACACAGGAGACTTAAAGTATGCTTTTACGGTTGAGATTCACCA	777		
QY	778	GTTTAACATTAAGGACAGAAATTTGATACCCCTACTATATATGAAACCGTGGTTTAC	795		
Db	796	GTTTAACATTAAGGACAGAAATTTGATACCCCTACTATATATGAAACCGTGGTTTAC	837		
QY	838	GATATCCCGAGGAGACAGAGCTTGTGAGCAATTTGAGATGACAGCAACATTCACAG	855		
Db	856	GATATCCCGAGGAGACAGAGCTTGTGAGCAATTTGAGATGACAGCAACATTCACAG	897		

QY	898	CCATAGGCCCACTCTTCTCTGAGCCGTGGTAAAGTAAATPPTGACAAAGCCGAAGCTGGG	957
Db	916	CCATAGGCCCACTCTTCTCTGAGCCGTGGTAAAGTAAATPPTGACAAAGCCGAAGCTGGG	975
QY	958	TCCCTGCTACTAAAGACAAGAAAGTTATCTGCTGAGCGATTCAGAGGCTTGAAGCCACCTAC	1012
Db	976	TCCCTGCTACTAAAGACAAGAAAGTTATCTGCTGAGCGATTCAGAGGCTTGAAGCCACCTAC	1033
QY	1018	TTCTTTCAGAAAAGCCCAATTCATATCTCACAGATCCAGCCCAATGGCGCTTGACAAAG	1077
Db	1036	TTCTTTCAGAAAAGCCCAATTCATATCTCACAGATCCAGCCCAATGGCGCTTGACAAAG	1099
QY	1078	TAAATGCTGACCGGGATAGCGAGGATGAAGTCGATGAGATGGTTCACAATTTTGAAGATC	1137
Db	1096	TAAATGCTGACCGGGATAGCGAGGATGAAGTCGATGAGATGGTTCACAATTTTGAAGATC	1155
QY	1138	GCCAGATGCTTGTGATGACTTTGTGTGATGTGAATTAAGATGAAGAAGCAATTCATGCATCTT	1197
Db	1156	GCCAGATGCTTGTGATGACTTTGTGTGATGTGAATTAAGATGAAGAAGCAATTCATGCATCTT	1218
QY	1198	GGAACTGCTGTTTGTAAAGAAACAAAGGGTTATACAGATGGTCATATCTCTTGGCGATGTG	1257
Db	1216	GGAACTGCTGTTTGTAAAGAAACAAAGGGTTATACAGATGGTCATATCTCTTGGCGATGTG	1279
QY	1258	AAGCATTTTTCAGAGATTTTACAGAGAAGAGTTGACACCGTTACTCATCATCTTCTGGTGT	1317
Db	1276	AAGCATTTTTCAGAGATTTTACAGAGAAGAGTTGACACCGTTACTCATCATCTTCTGGTGT	1339
QY	1318	GGAGATTTGTTTGTATTAACATATGAAACCAATGACCTTGTGACTCAGCCACCATCAACA	1377
Db	1336	GGAGATTTGTTTGTATTAACATATGAAACCAATGACCTTGTGACTCAGCCACCATCAACA	1399
QY	1378	ACATGCAATPACCATCTCGAGAAATTCGCGTATTAAGTTCAGACCAACCAACCAACCAACA	1437
Db	1396	ACATGCAATPACCATCTCGAGAAATTCGCGTATTAAGTTCAGACCAACCAACCAACCAACA	1459
QY	1438	ACAACAGTGTGGATCGTCCAGTCACTCAAAACACCAACCAATTAATTTGTGATCATC	1497
Db	1456	ACAACAGTGTGGATCGTCCAGTCACTCAAAACACCAACCAATTAATTTGTGATCATC	1519
QY	1498	CCAATGCAATTAACACACAGACAATGTGTCACAAACAAAGGACAAATTAACGACAGACAAG	1557
Db	1516	CCAATGCAATTAACACACAGACAATGTGTCACAAACAAAGGACAAATTAACGACAGACAAG	1579
QY	1558	TAAATTAATAGGAAATATCTCCGCTTTATATATACCATTTATGGATTTGAATCTATTC	1617
Db	1576	TAAATTAATAGGAAATATCTCCGCTTTATATATACCATTTATGGATTTGAATCTATTC	1639
QY	1618	TTCTTTCCTTAAAAAATTTGTATAGAGACAACAAATTTTATATGTATAGTATTAATCACT	1677
Db	1636	TTCTTTCCTTAAAAAATTTGTATAGAGACAACAAATTTTATATGTATAGTATTAATCACT	1699
QY	1678	GATTACATTTTGTAGTAAAAAATAA 1704	
Db	1696	GATTACATTTTGTAGTAAAAAATAA 1722	
RESULT 6			
LOCUS	AX032818	1715 bp	DNA
DEFINITION	Sequence 58 from Patent WO046358.	linear	PAT 21-SEP-2000
ACCESSION	AX032818		
VERSION	AX032818.1		
KEYWORDS	GI:10279794		
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	artificial sequences.		
AUTHORS	1 (bases 1 to 1715)		
TITLE	Dean, C., West, J. and Johanson, U.		
JOURNAL	Plant gene		
	Patent: WO 0046358-A 58 10-AUG-2000;		
	DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;		

JOHANSON URBAN (SE)									
FEATURES	Location	Offsets							
SOURCE	1..1715								
	/organism="synthetic_construct" /db_xref="taxon:32630" /note="Columbia VRN2 cDNA"								
BASE COUNT	536 a	326 c	353 g	500 t					
ORIGIN									
Query Match	91.1%;	Score 1569.6;	DB 6;	Length 1715;					
Best Local Similarity	96.5%;	Pred. No. 0;							
Matches 1649;	Conservative 0;	Missmatches 49;	Indels 10;	Gaps 4					
OY	1	CAAGTCTTCAATTTCGCTG--CTGCTCTTACACGCCAATGGGTGTTCCAGCT	58						
Db	1	CAAGTCTTCAATTTCGCTGCTCTCTCTTACACGCCCAATGGGTGTTCCAGCT	60						
OY	59	TTCAGGCCCTCAATCCAGACATTTCTATATTAAGCATTTGCAAGAGCGGTTCTAATTG	118						
Db	61	TTCAGGCCCTCAATACAGACATTTCTATATTAAGCATTTGCAAGAGCGGTTCTAATTG	120						
OY	119	TTGCATTGATTTTATCGCTATGACGTAGGGAATTTCTAATTTAGGGAGCGCTCAGAGTT	178						
Db	121	TTGCATTGATGTAACCAATATGACGTAGGGAATTTCTAATTTAGGGAGCGCTCAGAGTT	180						
OY	179	TGCATACCTTCATATGCGGCTCTGACGGTTGTAGTATTTGACAAGAAATGTGTG	238						
Db	181	TGCATACCTTCATATATGACGTCTGAGCGTTGTATTTGATTTGACAAGAAATGTGTG	240						
OY	239	GCAGAAATGTGCGCGGAATCCTCACCGGAGGAAGTGAATTCACGTGAGAAATCTTT	298						
Db	241	GCAGAAATGTGCGCGGAATCCTCACCGGAGGAAGTGAATTCACGTGAGAAATCTTT	300						
OY	299	GATATATTGTAAACCTGTTGCACATATATAACATCTTTACCTTCGCTCTGTAGCAACC	358						
Db	301	GATATATTGTAAACCTGTTGCACATATATAACATCTTTACCTTCGCTCTGTAGCAACC	360						
OY	359	ATCGTTCTCCCAAGATGCTTGAATACAAAATTGGACCAAGCCAAAGAAAGATCAAG	418						
Db	361	ATCGTTCTCCCAAGATGCTTGAATACAAAATTGGGCAAGCCCAAGAAAGATCAAG	420						
OY	419	ATCTACTGGAGTGTAGTTTTCACATATPAAGATGTATAACACATTACAGAAAACCTGA	478						
Db	421	ATCTACTGGAGTGTAGTTTTCACATATPAAGATGTATAATATACATTACAAAAGACTGA	480						
OY	479	AGTTAGGAGAGATGTCTGTGCCATTTTGCCTATAGCTATGTGTAGCTTCAAGGGCT	538						
Db	481	AGTTAGGAGAGATGTCTGTGCCATTTTGTCTATAGTATGTGTAGCTTCAAGGGGT	540						
OY	539	GCAATTTCAATTGAATTCATCATCATGATTTATTTAATTTGAGTTCACAGCTTTGAGAGA	598						
Db	541	GCAATTTCAATTGAATTCATCATCATGATTTATTTAATTTGAGTTCACAGCTTTGAGAGA	600						
OY	599	ATACACAGACGTTAATGTTCTGTAAACCTTAATTCCTCATATTTTGAGAGAAAGGAG	658						
Db	601	ATACACAGACGTTAATGTTCTGTAAACCTTAATTCCTCATATTTTGAGAGAAAGGAG	660						
OY	659	TGATGAGAGATAAATTTGAGCCCTTCTCTCTGCTGCTGCAAACTCGTAAAGCGAGACAAG	718						
Db	661	TGATGAGAGATAAATTTGAGCCCTTCTCTCTGCTGCTGCAAACTCGTAAAGCGAGACAAG	720						
OY	719	AGGTGGCGAATATAACACAGAGACTTAAAGTATGTTTTTACCGTTGGATTCAACCGAG	778						
Db	721	AGGTGGCGAATATAACACAGAGACTTAAAGTATGTTTTTACCGTTGGATTCAACCGAG	780						
OY	779	TTTAACTAATGGCAGAGAAAATGGAATCACCCCTACTTAATGATGAGAAACCGTGTTTAGG	838						
Db	781	TTTAACTAATGGCAGAGAAAATGGAATGACCCCTGCTGAATGATGAGAAACCGGTGTTAGG	840						
OY	839	ATAATCCGAGGCAACAGAGCTTGCTGGACATTTTGAGATGACACACATTCACACAGC	898						
Db	841	ATAATCCGAGGCAACAGAGCTTGCTGGACATTTTGAGATGACATTCACACATTCACACAGC	900						

QY 899 CATAGCCCACTCTTCTGACGCTGCTAAAGTTATTTGACACGAGAGCTGTGT 958
 DB 901 CATAGCCCACTCTTCTGACGCTGCTAAAGTTATTTGACACGAGAGCTGTGT 960
 QY 959 CCGGCTACTTAAGACAGAAAGTTATCTGAGGATCAAGAGCTTAAGCCACTACT 1018
 DB 961 CCGGCTACTTAAGACAGAAAGTTATCTGAGGATCAAGAGCTTAAGCCACTACT 1020
 QY 1019 TCTTCAGAAAGCCCAATCTATCATCTCCACAGACTCCAGCAATGGCGCTTGACAACT 1078
 DB 1021 TCTTCAGAAAGCCCAATCTATCATCTCCACAGACTCCAGCAATGGCGCTTGACAACT 1080
 QY 1079 AATGTCTGACCGGATAGGAGATGTAAGTCATGACGATGTTCGAGATTTTGAAAGATCG 1138
 DB 1081 AATGTCTGATCGGATAGGAGATGTAAGTCATGACGATGTTCGAGATTTTGAAAGATCG 1140
 QY 1139 CCAAGATCCTGATGACTTGTGTGATGTGAATTAAGTGAAGAAACCAATTCATCTTTG 1198
 DB 1141 CCAAGATCCTGATGACTTGTGTGATGTGAATTAAGTGAAGAAACCAATTCATCTTTG 1200
 QY 1199 GAACTCGTTGTGAAGAAACAAAGGTTATACAGATGTCATATCTCTGGGATGTGA 1258
 DB 1201 GAACTCGTTGTGAAGAAACAAAGGTTATACAGATGTCATATCTCTGGGATGTGA 1260
 QY 1259 AGCATTTTCAAGATTTTACAGAAAGTTGACCGCTTACATCATCTCTCTGCTGTG 1318
 DB 1261 AGTATTTTCAAGATTTTACAGAAAGTTGACCGCTTACATCATCTCTCTGCTGTG 1320
 QY 1319 GAGATTTTGTGATTAATTAATCTGAAACCATGAGCTTGTGACCTGACGACCATCAACAA 1378
 DB 1321 GAGATTTTGTGATTAATTAATCTGAAACCATGAGCTTGTGACCTGACGACCATCAACAA 1380
 QY 1379 CTGCAATACCATCTCTGAGAAATTCGCTAATAGTCAGACACACACACACACAA 1438
 DB 1381 CTGCAATACCATCTCTGAGAAATTCGCTAATAGTCAGACACACACACACACAA 1434
 QY 1439 CAACAGTGTGATGCTGCCAGTACTCAACACCAACCAATTAATCTGTGATCAATCC 1498
 DB 1435 CAACAGTGTGATGCTGCCAGTACTCAACACCAACCAATTAATCTGTGATCAATCC 1494
 QY 1499 CAATGACATTAACACAGACAGACATTTGACACAGACAAATTAACAGACAGAAAT 1558
 DB 1495 GAATGACATTAACACAGACAGACATTTGACACAGACAAATTAACAGACAGAAAT 1553
 QY 1559 AATTAATAGAAATCTCGGCTTTTATGATACCGATTTATCGGATTTGAATTAATCT 1618
 DB 1554 AATTAATAGAAATCTCGGCTTTTATGATACCGATTTATCGGATTTGAATTAATCT 1613
 QY 1619 TCTTCTCT-AAAAAATGTTAGAGCAACAAATTTTATATGTTAGTATTAATCACT 1677
 DB 1614 TCTTCTCTAAAAAATGTTAGAGCAACAAATTTTATATGTTAGTATTAATCACT 1673
 QY 1678 GATTACATTTTGTAAAAAATTAAT 1705
 DB 1674 GATTACATTTTGTAAAAAATTAAT 1701
 RESULT 7
 AX032891 1715 bp DNA linear PAT 21-SEP-2000
 LOCUS Sequence 4 from Patent WO004918.
 DEFINITION AX032891
 ACCESSION AX032891 GI:10279827
 VERSION
 KEYWORDS
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS Dean, C. and Gendall, A.
 TITLE Methods and means for modification of plant characteristics using the vernalization gene vrn2

JOURNAL Patent: WO 0044918-A 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD (GB)
 FEATURES
 source location/Qualifiers
 1. 1715
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 536 a 326 c 533 g 500 t
 ORIGIN
 Query Match 91.1%; Score 1569.6; DB 6; Length 1715;
 Best Local Similarity 96.5%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;
 1 CAAGCTTCTCAATTTTGGCTTG--CTCTCTTACACAGCAATCGGTTTTCACACT 58
 1 CAAGCTTCTCAATTTTGGCTTGCTCTCTCTTACACAGCAATCGGTTTTCACACT 60
 59 TTCAGCCCTCAATCCAGACATTTCTATATAGCATATTGACAGAGCGGTTCTAATG 118
 61 TTCAGCCCTCAATCCAGACATTTCTATATAGCATATTGACAGAGCGGTTCTAATG 120
 119 TTGCATGAGTTTATCGCTATGAGCTAGGGAATTTCTAATTTAGGGAGCGCTCAGACT 178
 121 TTGCATGAGTTTATCAAAATATGACCTAGGGAATTTCTAATTTAGGGAGCGCTCAGACT 180
 179 TTGCATGAGTTTATCAATCGCTCTGACGTTGTGATGTAATTTAGCAAGAAATGTGAG 238
 181 TTGCATGAGTTTATCAATCGCTCTGACGTTGTGATGTAATTTAGCAAGAAATGTGAG 240
 239 GAGAAATGTGCGCGGAAATTCACCGGAGAGAAATTTCAACGATGAAATCTCTT 298
 241 GAGAAATGTGCGCGGAAATTCACCGGAGAGAAATTTCAACGATGAAATCTCTT 300
 299 GATATATTTGAACCTGTTGACATATTAATCAATCTTCACTTCTGCTCTAGCAACCC 358
 301 GATATATTTGAACCTGTTGACATATTAATCAATCTTCACTTCTGCTCTAGCAACCC 360
 359 ATCTGTTCTTCAAGATGCTTGAACCTAACAATTTGAGCAAGCGCAAAAGAAATCAAG 418
 361 ATCTGTTCTTCAAGATGCTTGAACCTAACAATTTGAGCAAGCGCAAAAGAAATCAAG 420
 419 ATCTGTTCTTCAAGATGCTTGAACCTAACAATTTGAGCAAGCGCAAAAGAAATCAAG 478
 421 ATCTGTTCTTCAAGATGCTTGAACCTAACAATTTGAGCAAGCGCAAAAGAAATCAAG 480
 479 AGTTAGGAGATGTTCTTCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 538
 481 AGTTAGGAGATGTTCTTCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 540
 539 GCAATTTCAATTTGATCATCTCATGATTTATTTGATTTGATTTGATTTGATTTGATTT 598
 541 GCAATTTCAATTTGATCATCTCATGATTTATTTGATTTGATTTGATTTGATTTGATTT 600
 599 ATACAGACATTAATGTTCTGTAATTAATTTCTTCAATTTCTTCAATTTCTTCAATTT 658
 601 ATACAGACATTAATGTTCTGTAATTAATTTCTTCAATTTCTTCAATTTCTTCAATTT 660
 659 TGATGACATTAATTTGAGCCCTCTCTCTCTCTGAAACCTGTAAGCGAGACAAAG 718
 661 TGATGATATTAATTTGAGCCCTCTCTCTCTCTGAAACCTGTAAGCGAGACAAAG 720
 719 AGTGGCGAAATTAACACAGAGACTTAAGATATGCTTTTACGTTGATTCACCGAG 778
 721 AGTGGCGAAATTAACACAGAGACTTAAGATATGCTTTTACGTTGATTCACCGAG 780
 779 TTAACTAATGACAGAGAAATGGAATCACTACTTAATGATGGAACCGGTGTTAGG 838
 781 TTAACTAATGACAGAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 840
 839 ATATCCGAGGACAGAGGCTGCTGACAAATTTGATGACGACGAAATTCACCGAGC 898
 841 ATATCCGAGGACAGAGGCTGCTGACAAATTTGATGATGATGATGATGATGATGATGAT 900

QY	899	CATAGCCCACTCTCTCTGGACGCTGGGCTCAAAGTATATTTGACAAACGGAAGCTGGGT	958
Db	901	CATAGCCCACTCTCTCTGGACGCTGGGCTCAAAGTATATTTGACAAACGGAAGCTGGGT	960
QY	959	CCCTGCTACTAAGACAAAGAAAGTTATCTGCTAGCGATCAGAGGCTAGAACCCACTACT	1018
Db	961	CCCTGCTACTAAGACAAAGAAAGTTATCTGCTAGCGATCAGAGGCTAGAACCCACTACT	1020
QY	1019	TCCTCAGAAAACCCCAATTTATCATCTTCCACAGAGTCCAGCCAAATGGCGCTTGAGCAAGT	1078
Db	1021	TCCTCAGAAAACCCCAATTTATCTATCTTCCACAGAGTCCAGCCAAATGGCGCTTGAGCAAGT	1080
QY	1079	AATGTCGACCGGGATGCGAGATGGAAGTCAGATGACGATGTGGAGATTTTGAAGATCG	1138
Db	1081	AATGTCGATCGGGATGCGAGATGGAAGTCAGATGACGATGTGGAGATTTTGAAGATCG	1140
QY	1139	CCAGATGCTTGATGACTTTGTGATGTGAATTAAGATGAAAAAGCAATTCATGCATCTTGG	1198
Db	1141	CCAGATGCTTGATGACTTTGTGATGTGAATTAAGATGAAAAAGCAATTCATGCATCTTGG	1200
QY	1199	GAACCTGCTTGTAAGAAAACAAAGGTTATAGCAGATGGTATATCTTGGGATGTGA	1255
Db	1201	GAACCTGCTTGTAAGAAAACAAAGGTTATAGCAGATGGTATATCTTGGGATGTGA	1260
QY	1259	AGCATTTTCAAGATTTTACGAGAAAGATGCGACCGTACTCATACACTCTTGGGTGTG	1318
Db	1261	AGATTTTCAAGATTTTACGAGAAAGATGCGACCGTACTCATACACTCTTGGGTGTG	1320
QY	1319	GAGATTTGTTTGTATTAACTAGTGAACCATGACTGTGACTGACGACCAATCAACAA	1378
Db	1321	GAGATTTGTTTGTATTAACTAGTGAACCATGACTGTGACTGACGACCAATCAACAA	1380
QY	1379	CTGCAATACCATCTCGAGAAATGGCCGTAATAGTCTGACACCAACCCACCAACAA	1438
Db	1381	CTGCAATACCATCTCGAGAAATGGCCGTAATAGTCTGACACCAACCCACCAACAA	1440
QY	1439	CAACAGTGTGGATCGTCCGAGTGACTCAAAACCAACCAATTAATTTGGATCATCC	1498
Db	1435	CAACAGTGTGGATCATCCGAGTGACTCAAAACCAACCAATTAATTTGGATCATCC	1499
QY	1499	CAATGACATTAACACAAAGAAACAATGTTGACACAGGACATTAACAGCAGACAAAGT	1558
Db	1495	GAATGACATTAACAAAGAAACAATGTTGACACAGGACATTAACAGCAGACAAAGT	1553
QY	1559	AATTAATTAAGAAAATCTCGGCTTTATGATACCGATTTATGCGATTTGTAACATTAATCT	1618
Db	1554	AATTAATTAAGAAAATCTCGGCTTTATGATACCGATTTATGCGATTTGTAACATTAATCT	1613
QY	1619	TCCTTTCTT-AAAAAATGTTTAGAGCAACAATTTTATGTTAGTATTAATCACT	1677
Db	1614	TCCTTTCTTAAAAAATGTTTAGAGCAACAAGATTTTATGTTAGTATTAATCACT	1673
QY	1678	GATTACATTTTGTAGTAAAAAAAAT 1705	
Db	1674	GATTACATTTTGTAGTAAAAAAAATGAT 1701	
RESULT 8			
AY034902			
LOCUS			
DEFINITION			
AY034902 1721 bp mRNA linear PLN 23 -APR-2002			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
Yamada, K., Liu, S.-X., Sakano, H., Pham, P. K., Banh, T., Chung, M. K.,			

TITLE JOURNAL REFERENCE
AUTHORS Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
 Arabidopsis full length cDNA clones
 unpublished
 2 (bases 1 to 1721)

TITLE JOURNAL REFERENCE
COMMENT Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES
source Location/Qualifiers
 1..1721
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="4"
 /clone="RAFL05-08-K19 (R09871)"
 /note="This clone is in a modified pluescript vector (Lambda ZAP) as a XhoI/SstI insert.
 ecotype: Columbia"
 1..1721
 /gene="At4g16845"
 1..231
 /gene="At4g16845"
 232..1554
 /gene="At4g16845"
 /codon_start=1
 /evidence=experimental
 /product="unknown protein"
 /protein_id="AAK59409.1"
 /db_xref="GI:14334422"
 /translation="MCRONCRAKSSPEEIVSTDENLLIYCKPRLYNIFHLRSGNPEFLPRLKIGAKRRKRKRSRSGVYFVYKCDNNLTLOREVEDSCSPGSLGSPFGKLOPHNSHDLFEFEKRLREYQVNVSYKNTALINDGNRLGIPEATELGQFEMTSROKGRNNTRLRKACFLPLSPSLANGENKILIPLESPLANSKISLIPLESPLANSKIPALPAHNSLIDAGAKVILTEAVAPATKTRKLSAERSSEASHLLRQYHSHRQVAPMLAEYVMSDRSDSEDDVDVADDEQMDLDEVDYVADDEKQFHLNLSFYRKQRYLADGHSWACVEFYSREKELICYSLIFWCWRLFLIKLNNHGLVDSATINCWTILENCRNSTGNTNNNSVHPDSDSNTNNNNIVDHPDIDKKNKNNVDKNNDSRDK"
 1555..1721
 /gene="At4g16845"
 1555..1721

BASE COUNT 548 a 324 c 353 g 496 t
ORIGIN

Query Match 91.1% Score 1568.6; DB 8; Length 1721;
 Best Local Similarity 96.5% Pred. No. 0;
 Matches 1648; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

OY 2 AACCTTCTCAATTTTGGCTTG--CTCTCTCTTACACAGCAATGGTGTTCGACCTT 59
 DB 1 AAGCTTCTCAATTTTGGCTTGCTCTCTCTTACACAGCAATGGTGTTCGACCTT 60
 OY 60 TCAGGCTCAATTCAGACATCTCTATATAGCATTTTCAGAGAGGGGGTCTTAATGT 119
 DB 61 TCAGGCTCAATTCAGACATCTCTATATAGCATTTTCAGAGAGGGGGTCTTAATGT 120
 OY 120 TGCATGAGTTTATGCTATGAGCTAGAGAAATCTATATAGGAGGCTCTGACCTT 179
 DB 121 TGCATGAGTTTATGAGCATATGAGCTAGAGAAATCTATATAGGAGGCTCTGACCTT 180
 OY 180 GCACCTACCTCATATGAGCTCTGAGCTGTGTATGTTGTAACAGAAATGTGAGG 239
 DB 181 GCACCTACCTCATATGAGCTCTGAGCTGTGTATGTTGTAACAGAAATGTGAGG 240
 OY 240 CAGAAATGTGCGGCAATTCCTACCGAGAGAAATGATTCAGTATGAGAAATCTCTG 299
 DB 241 CAGAAATGTGCGGCAATTCCTACCGAGAGAAATGATTCAGTATGAGAAATCTCTG 300
 OY 300 ATATATTTGTAACCTGCTTCTGACTATATACATCTTTCACCTTCTCTGAGCAACCA 359
 DB 301 ATATATTTGTAACCTGCTTCTGACTATATACATCTTTCACCTTCTCTGAGCAACCA 360
 OY 360 TCGTTTCTTCCAAAGATGCTTGAATACAAATTTGGAGCAAGCGCAAAAGAAATGTCA 419
 DB 361 TCGTTTCTTCCAAAGATGCTTGAATACAAATTTGGAGCAAGCGCAAAAGAAATGTCA 420
 OY 420 TCTACTGGATGATGATTTTCAACTATTAAGATTTGATTAACATTAACAGAAATCTGAA 479
 DB 421 TCTACTGGATGATGATTTTCAACTATTAAGATTTGATTAACATTAACAGAAATCTGAA 480
 OY 480 GTTACGAGAGATTTCTTCTGATTTTCTCTATATGATGATGATGATGATGATGATG 539
 DB 481 GTTACGAGAGATTTCTTCTGATTTTCTCTATGATGATGATGATGATGATGATGATG 540
 OY 540 CAATTTCAATTTGATTCATCTCATATGATTTTGAATTTGATTTGATTTGATTTGATTT 599
 DB 541 CAATTTCAATTTGATTCATCTCATATGATTTTGAATTTGATTTGATTTGATTTGATTT 600
 OY 600 TACACAGAGATTAATGTTTCTGTAAACTTAATTCCTTCAATTTTGAAGAAAGAGAT 659
 DB 601 TACACAGAGATTAATGTTTCTGTAAACTTAATTCCTTCAATTTTGAAGAAAGAGAT 660
 OY 660 GATGAGATTAATTTGAGCCCTCTCTCTCTCTCTGGAACCTCGTAAGCGAGACAAGA 719
 DB 661 GATGAGATTAATTTGAGCCCTCTCTCTCTCTCTGGAACCTCGTAAGCGAGACAAGA 720
 OY 720 GGTGGAGAAATTAACCCGAGAGACTTAAGATGATTTTACCGTTGATTTACCCAGT 779
 DB 721 GGTGGAGAAATTAACCCGAGAGACTTAAGATGATTTTACCGTTGATTTACCCAGT 780
 OY 780 TTACTTAATGCGACAGAAATGAAATCAACCTACTTAATGATGAAACCGTGTGTAAGA 839
 DB 781 TTACTTAATGCGACAGAAATGAAATCAACCTACTTAATGATGAAACCGTGTGTAAGA 840
 OY 840 TATCCGAGAGCAACAGAGCTTCTGAGCAATTTGAGATGACCAACATTTCCACAGCC 899
 DB 841 TATCCGAGAGCAACAGAGCTTCTGAGCAATTTGAGATGACCAACATTTCCACAGCC 900
 OY 900 ATAGCCCACTTCTCTGAGAGCTGCTGCTAAAGTTATGATGAGCAAGGAAAGCTGTGTC 959
 DB 901 ATAGCCCACTTCTCTGAGAGCTGCTGCTAAAGTTATGATGAGCAAGGAAAGCTGTGTC 960
 OY 960 CCGGCTACTAAGCAAGAAAGTATCTGCTGAGGATGAGGCTAAGAGCCACTTACTT 1019
 DB 961 CCGGCTACTAAGCAAGAAAGTATCTGCTGAGGATGAGGCTAAGAGCCACTTACTT 1020
 OY 1020 CTTTCAGAAAGCCAAATTTATCATTTCTACAGAGTCCAGCCAAATGGCGCTTGACAAAGTA 1079
 DB 1021 CTTTCAGAAAGCCAAATTTATCATTTCTACAGAGTCCAGCCAAATGGCGCTTGACAAAGTA 1080

OY 1080 ATGCTGACCGGATAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
 DB 1081 ATGCTGATCGGATAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 OY 1140 CAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
 DB 1141 CAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 OY 1200 AACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
 DB 1201 AACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 OY 1260 GCATTTTCAAGATTTTACAGAAAGATTTCCACCTTACTCATCTCTTCTGCTGTGG 1319
 DB 1261 GCATTTTCAAGATTTTACAGAAAGATTTCCACCTTACTCATCTCTTCTGCTGTGG 1320
 OY 1320 AGATTTGTTTGTGATTAACATTAAGCAATGAGATGATGATGATGATGATGATGATGATG 1379
 DB 1321 AGATTTGTTTGTGATTAACATTAAGCAATGAGATGATGATGATGATGATGATGATGATG 1380
 OY 1380 TGCATTAACATCTCTGAGAAATTTGCGTAATAGCTCAGACACACACACACACACACAC 1439
 DB 1381 TGCATTAACATCTCTGAGAAATTTGCGTAATAGCTCAG-----TCACTAACACACACAC 1434
 OY 1440 AACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1499
 DB 1435 AACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1494
 OY 1500 AATGACATTAAC 1559
 DB 1495 AATGACATTAAC 1553
 OY 1560 AATTAATGAGAAATCTCCGCTTTATGATTAACATTTATGATGATGATGATGATGATG 1619
 DB 1554 AATTAATGAGAAATCTCCGCTTTATGATTAACATTTATGATGATGATGATGATGATG 1613
 OY 1620 CTTTCTT-AAAAAATGTTTGAAGCAACAAATTTTATATGATGATGATGATGATGATG 1678
 DB 1614 CTTTCTT-AAAAAATGTTTGAAGCAACAAATTTTATGATGATGATGATGATGATGATG 1673
 OY 1679 ATTACATTTTATGATTAACAAATTAAT 1705
 DB 1674 ATTACATTTTATGATTAACAAATTAAT 1700
 RESULT 9
 AX032894 1737 bp DNA linear PART 21-SEP-2000
 LOCUS
 DEFINITION
 ACCESSION AX032894
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 1737)
 AUTHORS
 Dean, C. and Gendall, A.
 TITLE
 Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 Patent: WO 0044918-A 7 03-AUG-2000;
 JOURNAL
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)
 FEATURES
 source location/Qualifiers
 1..1737
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 543 a 330 c 359 g 505 t
 ORIGIN
 Query Match 89.3%; Score 1537.6; DB 6; Length 1737;
 Best Local Similarity 95.3%; Pred. No. 0;

Matches 1649; Conservative 0; Mismatches 49; Indels 32; Gaps 5;

```

OY 1 CAAGCTTCTCAATTTGCTG--CTCTCTTACACAGCAATGCGTCTTTCGAGCT 58
Db 1 CAAGCTTCTCAATTTGCTGCTCTCTCTTACACGCGCAATCGTCTTTCGAGCT 60
OY 59 TTCAGGCTTCATTCAGACATTTATATAGCATATTCAGAGAGCGGCTTCAATG 118
Db 61 TTCAGGCTTCATTCAGACATTTATATAGCATATTCAGAGAGCGGCTTCAATG 120
OY 119 TTGCATGAGTTTATGCTATGACGTAGGAAATTTCTAATTAAGGCGGCTCAGATT 178
Db 121 TTGCATGAGTTTATGCAATATATGACGTAGGAAATTTCTAATTAAGGCGGCTCAGAGTT 180
OY 179 TGCATTAATCTATATGCGCTCTGAGCTTGTGAGTATTAAGCAAGAAATGCTG 238
Db 181 TGCATTAATCTATATGCGCTCTGAGCTTGTGAGTATTAATTAAGCAAGAAATGCTG 240
OY 239 GCAGAAATGTCGCGGAAATCCTCAACGAGAGAGATTTGCACTGATGAAATCTTT 298
Db 241 GCAGAAATGTCGCGGAAATCCTCAACGAGAGAGATTTGCACTGATGAAATCTTT 300
OY 299 GATATATGTAACCTGTTGCACTATATATACATTTTCACTTCTCTAGGCAACC 358
Db 301 GATATATGTAACCTGTTGCACTATATATACATTTTCACTTCTCTAGGCAACC 360
OY 359 ATCGTTTCTCCAAAGTGGTGAATACAAATTTGGAGCAAGCGCAAGAAAGCAAG 418
Db 361 ATCGTTTCTCCAAAGTGGTGAATACAAATTTGGAGCAAGCGCAAGAAAGCAAG 420
OY 419 ATCTACTGAGATGATTTTCACTATATAGATTTGTAATTAACATTTACAGAAATCTGA 478
Db 421 ATCTACTGAGATGATTTTCACTATATAGATTTGTAATTAACATTTACAGAAATCTGA 480
OY 479 ACTTGGAGAGATTTGTTCTTCTGCAATTTGCTCTATGCTATGCTGAGCTTCA----- 531
Db 481 ACTTGGAGAGATTTGTTCTTCTGCAATTTGCTCTATGCTATGCTGAGCTTCAAGGCTG 540
OY 532 -----AGGCGCTGCAATTTGTAATCATCTCATGATTTATTTGAT 576
Db 541 CAACATTTACAACTAGAGGCTGCAATTTGTAATCATCTCATGATTTATTTGAT 600
OY 577 TTGAGTTCAAGCTTTTCGAAAGATACAGACAGTTAATGTTTCTGTAAGTAATTCCT 636
Db 601 TTGAGTTCAAGCTTTTCGAAAGATACAGACAGTTAATGTTTCTGTAAGTAATTCCT 660
OY 637 TCATATTTGAGAAAGAGATGATGACGATTAATTTGAGCCCTTCTCTCTGCTCGA 696
Db 661 TCATATTTGAGAAAGAGATGATGATGATTAATTTGAGCCCTTCTCTCTGCTCGA 720
OY 697 AACCTGTAAGCGAGACAAAGAGTGGCAGAAATTAACCGAGAGACTTAAGTATGCT 756
Db 721 AACCTGTAAGCGTATGACAAAGAGTGGCAGAAATTAACCGAGAGACTTAAGTATGCT 780
OY 757 TTTTACCGTTGATTTACCCAGTTTAACTAATGACAGAAATGGAATCACCTACTTA 816
Db 781 TTTTACCGTTGATTTACCCAGTTTAACTAATGACAGAAATGGAATGCTCTGCTGA 840
OY 817 ATGATGGAACCGTGTGTTAGATATCCCGAGCAACAGAGCTTGTGACAAATTTGACA 876
Db 841 ATGATGGAACCGTGTGTTAGATATCCCGAGCAACAGAGCTTGTGACAAATTTGACA 900
OY 877 TGACAGCAACATTTCCACAGGCAATAGCCACTTCTCTGAGCGTGTGTAAGTTA 936
Db 901 TGACAGCAACATTTCCACAGGCAATAGCCACTTCTCTGAGCGTGTGTAAGTTA 960
OY 937 TATTGACAGCAAGACTGTGCTCTGCTACTAATGACAGAAAGTTATCTGCTGAGCAT 996
Db 961 TATTGACAGCAAGACTGTGCTCTGCTACTAATGACAGAAAGTTATCTGCTGAGCAT 1020
OY 997 CAGAGGCTAAGAGCCACTACTTCTTCAAGAAAGCCCAATTTCTATCTTCTCACAGAGTCC 1056
Db 1021 CAGAGGCTAAGAGCCACTACTTCTTCAAGAAAGCCCAATTTCTATCTTCTCACAGAGTCC 1080

```

```

OY 1057 AGCCAAATGGCGCTTGAGCAAGTAATGCTGACCGGATAGCCAGATGAAGTCAGAC 1116
Db 1081 AGCCAAATGGCGCTTGAGCAAGTAATGCTGATCGGATAGCCAGATGAAGTCAGAC 1140
OY 1117 ATGTTGCAATTTTGAAGATTCGCGAGATGCTGATGACTTGTGATGTAATGAATG 1176
Db 1141 ATGTTGCAATTTTGAAGATTCGCGAGATGCTGATGACTTGTGATGTAATGAATG 1200
OY 1177 AAAGCAATTCATGCAATCTTTGGAAGTCTTTGTAAAGAAACAAGGCTTATAGCATG 1236
Db 1201 AAAGCAATTCATGCAATCTTTGGAAGTCTTTGTAAAGAAACAAGGCTTATAGCATG 1260
OY 1237 GTCATATCTCTTGGGATGTAAGATTTTCAAGATTTTCAAGAAAGGTTGCAAGCT 1296
Db 1261 GTCATATCTCTTGGGATGTAAGATTTTCAAGATTTTCAAGAAAGGTTGCAAGCT 1320
OY 1297 ACTCATCACTCTTCTGCTGATGAGATTTTGTGATTAAGTATGAAACATGAGACTTG 1356
Db 1321 ACTCATCACTCTTCTGCTGATGAGATTTTGTGATTAAGTATGAAACATGAGACTTG 1380
OY 1357 TCGACTGACCCACATCAACATGCAATACCATCTCGAGATTTGCCGTAATAGCTCAG 1416
Db 1381 TCGACTGACCCACATCAACATGCAATACCATCTCGAGATTTGCCGTAATAGCTCAG 1440
OY 1417 ACACGACACACACACACACACACACACACAGTGTGATGCTCCAGTCAACACACACA 1476
Db 1441 -----TCCTATACAAACACACACACACAGTGTGATGCTCCAGTCAACACACACA 1494
OY 1477 ACAATTAACATTTGTGATCATCCCAATGACATTAACCAAGAAACAATGTTGACAAACAG 1536
Db 1495 ACAATTAACATTTGTGATCATCCCAATGACATTAACCAAGAAACAATGTTGACAAACAG 1554
OY 1537 ACAATTAACGACAGACAAAGTAATTAATAGAAATATCCGCTTTTATGATACCGAT 1596
Db 1555 ACAATTAACGACAGAGAC- AAGTAATTAATAGAAACACCTCGGTTAATGATATACGAT 1613
OY 1597 TTATGAGTTGATCACTATTTCTCTTCTT-AAAAAATGTTAGAGCAACAATTTT 1655
Db 1614 CTATCGAGTTGATCACTATTTCTCTTCTTAAAAAATGTTAGAGCAACAAGATTT 1673
OY 1656 TTATATGTTAGTATCACTGATTAATTAATTTTAAAAAATTT 1705
Db 1674 TATTTTGTAGTATCACTGATTAATTTTAAAAAATTTAGT 1723

RESULT 10
AY063047 1354 bp mRNA linear PLN 18-APR-2002
LOCUS Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete
DEFINITION
cds.
ACCESSION AY063047
VERSION AY063047.1 GI:17104664
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
REFERENCE 1 (bases 1 to 1354)
Yamada, K., Liu, S. X., Sakano, H., Pham, P. K., Banh, J., Etgu, P.,
Lee, J. M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H.,
Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M.,
Palm, C. J., Shinn, P., Southwick, A., Davis, R. W., Ecker, J. R. and
Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1354)
Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J. M., Goldsmith, A. D.,
Lee, J. M., Onodera, C. S., Quech, H. L., Tang, C. C., Toriumi, M., Wu, H. C.,
Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,
Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A.,
Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Liu, J.,

```

TITLE
JOURNAL

COMMENT

Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A., Direct Submission
Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLP CDNA's (RFLP cDNA : RIKEN Arabidopsis Full-length cDNA) : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RFLP cDNAs: Yamada, K., Bann, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PIs.

FEATURES

source

1. 1354
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U09871"
/note="this clone is in PUNI 51.
ecotype: Columbia"
1. 1354
/gene="At4g16845"
1. 1323
/gene="At4g16845"
/codon_start=1
/evidence="experimental
/product="unknown protein"
/protein_id="AA134221.1"
/db_xref="GI:1710465"
/translation="MCRQNCRAKSSPEVYSTDENLLIYCKFVRLXNIFHLRLNPS
FLPRLNLTGKRRKRSSTGTVFNKYKCNNTLQRTVREDSCSPCSMLCGSFK
LQFLNNSHDLFEERKLEEGYIVANVSKNFTLEEDSDDEKPEPSLCSKRGK
RQGRNTRRLKYCFPLDSPSLANGTEGLALNDGRLGLPEATELAGEFMRTS
NIPPAIHSSLDAGAKVILTEAVVPATKRLSARSBSRSHLLQKROFVSHRYO
PMALQVNSDRSDSEVDVADPEROMLDVFNKDEKQFMHLMNSFVKQKQVIA
DGHISWACEVSRFEKELHCYSLEFCWMLFLIKLNHGLKQDSATINNCTILENCR
NTSYTNNNNNSVDHPSDSNTNNNNIYDHPIDIKNNKNNVDKNKNSRDK"

gene
CDS

1. 1323
/gene="At4g16845"
/codon_start=1
/evidence="experimental
/product="unknown protein"
/protein_id="AA134221.1"
/db_xref="GI:1710465"
/translation="MCRQNCRAKSSPEVYSTDENLLIYCKFVRLXNIFHLRLNPS
FLPRLNLTGKRRKRSSTGTVFNKYKCNNTLQRTVREDSCSPCSMLCGSFK
LQFLNNSHDLFEERKLEEGYIVANVSKNFTLEEDSDDEKPEPSLCSKRGK
RQGRNTRRLKYCFPLDSPSLANGTEGLALNDGRLGLPEATELAGEFMRTS
NIPPAIHSSLDAGAKVILTEAVVPATKRLSARSBSRSHLLQKROFVSHRYO
PMALQVNSDRSDSEVDVADPEROMLDVFNKDEKQFMHLMNSFVKQKQVIA
DGHISWACEVSRFEKELHCYSLEFCWMLFLIKLNHGLKQDSATINNCTILENCR
NTSYTNNNNNSVDHPSDSNTNNNNIYDHPIDIKNNKNNVDKNKNSRDK"

3'UTR
BASE COUNT
ORIGIN

431 a 264 c 287 g 372 t

Query Match 74.2% Score 1277.4; DB 8; Length 1354;
Best local similarity 97.2%; Pseq No. 2.4e-272;
Matches 1333; Conservative 0; Mismatches 31; Indels 7; Gaps 2;

QY 231 ATGTGTAGGAGATTTGGCGGCAAAATCTCACCGAGAGAGATTTCAACTGATGAG 290
DB 1 ATGTGTAGGAGAAATGTGCGGCAAAATCTCACCGAGAGAGATTTCAACTGATGAG 60
QY 291 AATCTCTGATATATGTAAACCTGTTCGACTATATAAATCTTTCACCTTGCTCTTA 350
DB 61 AATCTCTGATATATGTAAACCTGTTCGACTATATAAATCTTTCACCTTGCTCTTA 120
QY 351 GGCAACCATGTTCTCCAAAGATGCTTGAACCTAACAAATTTGAGCAAGGCAAAAGA 410
DB 121 GGCAACCATGTTCTCCAAAGATGCTTGAACCTAACAAATTTGAGGCAAGGCAAAAGA 180
QY 411 AAGTCAAGTCTACGGGATGAGTGTTCACCTATTAAGATGATGATTAACCATTAAG 470
DB 181 AAGTCAAGTCTACGGGATGAGTGTTCACCTATTAAGATGATGATTAACCATTAAG 240

QY 471 AAACGTAAGTATAGGAGAGATTTGCTTCCATTTTGCCTATAGTATGTGATGCTTC 530
DB 241 AGACGTAAGTATAGGAGAGATTTGCTTCCATTTTGCCTATAGTATGTGATGCTTC 300
QY 531 AAGGCGTCAATTTCAATTTCAATTCATTCATGATTTATTTGAATTTGATTTCAAGCTT 590
DB 301 AAGGCGTCAATTTCAATTTCAATTCATTCATGATTTATTTGAATTTGATTTCAAGCTT 360
QY 591 TTGGAAGATTTACAGACGTTAATGTTCTGTAAACCTTAATTCCTTCAATTTAGAGAA 650
DB 361 TTGGAAGATTTACAGACGTTAATGTTCTGTAAACCTTAATTCCTTCAATTTAGAGAA 420
QY 651 GAAGGAAGTATAGGAGATTTGAGCCCTTCTCTCTGCTGCAAACTGTAAGCG 710
DB 421 GAAGGAAGTATAGGAGATTTGAGCCCTTCTCTCTGCTGCAAACTGTAAGCGT 480
QY 711 AGCAAAAGAGTGGAGAAATTAACACGAGAGACTTAAGTATGCTTTTACCGTTGAT 770
DB 481 AGCAAAAGAGTGGAGAAATTAACACGAGAGACTTAAGTATGCTTTTACCGTTGAT 540
QY 771 TCACCAAGTTTAACTAATGACAGAAAATGGAATCACCTTAATGATGAAACCGT 830
DB 541 TCACCAAGTTTAACTAATGACAGAAAATGGAATGCAATTTGATGATGAAACCGT 600
QY 831 GGTATAGATATCCGAGGCAACAGAGCTTGTGACAAATTTGATGACCAACATTT 890
DB 601 GGTATAGATATCCGAGGCAACAGAGCTTGTGACAAATTTGATGACCAACATTT 660
QY 891 CCACAGCATATGCCACTCTTCTCTGAGCGTGTGCTAAATTTATTTGCAAGCGAA 950
DB 661 CCACAGCATATGCCACTCTTCTCTGAGCGTGTGCTAAATTTATTTAACAACCGAA 720
QY 951 GCGTGTGCTCCGCTGCTAATGAGAAAGTATCTGCTGAGGATGACAGGCTGAAGC 1010
DB 721 GCGTGTGCTCCGCTGCTAATGAGAAAGTATCTGCTGAGGATGACAGGCTGAAGC 780
QY 1011 CACCTACTTCTTCAAAAAGCCCAATTCATCTCTCAAGATCCACCAATGGCGCTT 1070
DB 781 CACCTACTTCTTCAAAAAGCCCAATTCATCTCTCAAGATCCACCAATGGCGCTT 840
QY 1071 GAGCAAGTATGTCTGACCGGAGATAGGAGATGAATGATGATGATGATGATGAT 1130
DB 841 GAGCAAGTATGTCTGATGAGGAGATAGGAGATGAATGATGATGATGATGATGAT 900
QY 1131 GAAGATCGCAGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1190
DB 901 GAAGATCGCAGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1191 CATCTTTGAGACTGCTTTGTAAGAAACAAAGGTTTATACAGATGCTATATCTCTGG 1250
DB 961 CATCTTTGAGACTGCTTTGTAAGAAACAAAGGTTTATACAGATGCTATATCTCTGG 1020
QY 1251 GCATGTGAAGCATTTTCAAGATTTTGAAGAAAGTGTGACCGTTATGATGATGATGAT 1310
DB 1021 GCATGTGAAGCATTTTCAAGATTTTGAAGAAAGTGTGACCGTTATGATGATGATGAT 1080
QY 1311 TGGTGTGAGATTTTGTGATTTAATTAATGAAGACCATGACTTGTGACTGACCCACC 1370
DB 1081 TGGTGTGAGATTTTGTGATTTAATTAATTAATGAAGACCATGACTTGTGACTGACCCACC 1140
QY 1371 ATCAACAACTGCAATACCATCTCGAAGAAATGCGGTAAATGCTGACACACCAACCC 1430
DB 1141 ATCAACAACTGCAATACCATCTCGAAGAAATGCGGTAAATGCTGACACACCAACCC 1194
QY 1431 AACCAACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490
DB 1195 AACCAACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
QY 1491 GATCAATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1550
DB 1255 GATCAATCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314

OY		1551	GACAAAGTAATTTAATFAGAAAATTCCTCGGCTTTTATGATA	1591
Db		1315	GAC-AAAGTAATTAAATAGAACACACTCCGGTTTAGATGATA	1354
RESULT_11				
LOCUS	AX032890		6338 bp	DNA linear PAT 21-SEP-2000
DEFINITION	Sequence 3 from Patent WO0044918.			
ACCESSION	AX032890			
VERSION	AX032890.1		GI:10279826	
KEYWORDS				
SOURCE				
ORGANISM				
	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
	1 (bases 1 to 6338)			
	Dean,C. and Gendall,A.			
	Methods and means for modification of plant characteristics using			
	the vernalization gene vrn2			
	Patent: WO 0044918-A 3 03-AUG-2000;			
	DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD			
JOURNAL	(GB)			
FEATURES				
source	location/Qualifiers			
	1..6338			
	/organism="Arabidopsis thaliana"			
	/db_xref="taxon:3702"			
BASE COUNT	1932 a 1076 c 1134 g 2193 t		3 others	
ORIGIN				
	Query Match	23.9%; Score 412.2; DB 6; Length 6338;		
	Best Local Similarity	98.8%; Pred. No. 8.7e-81;		
	Matches 414; Conservative	1; Mismatches 4; Indels 0; Gaps 0;		
OY	1304	ACTCTTGCGGTGGAGATTGTTTGGATTGAACAATAAGAGACCATGGACTTGCAGCTC	1363	
Db	5102	AATTTCCAGGTGGTGGAGATTGTTTGGATTGAACAATAAGAGACCATGGACTTGCAGCTC	5161	
OY	1364	AGGCCACCATCAACAACTGCAATACACTCTCGAGAAATTCGGTAATAGCTCAGACACCAC	1423	
Db	5162	AGGCACCATCAACAACTGCAATACACTCTCGAGAAATTCGGTAATAGCTCAGACACCAC	5221	
OY	1424	CACCACCACAAACAACAACAGTGTGATCGTCCAGTACTCAACAACCAACAATPAA	1483	
Db	5222	CACCACCACAAACAACAACAGTGTGATCGTCCAGTACTCAACAACCAACAATPAA	5281	
OY	1484	CATTGTGATCATCCCAATGACATGAACAACAACAAGAATGTGACAAACAAGCAATPAA	1543	
Db	5282	CATTGTGATCATCCCAATGACATGAACAACAACAAGAATGTGACAAACAAGCAATPAA	5341	
OY	1544	CAGCAGAGACAAGATAATTAATAGAAAAATCTCCGCTTTTATGATACGATTTATCGG	1603	
Db	5342	CAGCAGAGACAAGATAATTAATAGAAAAATCTCCGCTTTTATGATACGATTTATCGG	5401	
OY	1604	ATTGTAACTAATTCCTCTTCTTAATAAAATTTGTTAGAGCAACAATTTTTATATGT	1663	
Db	5402	ATTGTAACTAATTCCTCTTCTTAATAAAATTTGTTAGAGCAACAATTTTTATATGT	5461	
OY	1664	TAGGTATTTGACATGATTACATTTTATGTTAAAAAATGATATCTGTTTAACT	1722	
Db	5462	TAGGTATTTGACATGATTACATTTTATGTTAAAAAATGATATCTGTTTAACT	5520	
RESULT_12				
LOCUS	ATRP5LE2/c			
DEFINITION	Arrip5le2	91660 bp	DNA linear	PLN 22-DEC-2000
ACCESSION	AF180942.1			
VERSION	AF180942.1		GI:6449044	
KEYWORDS				
SEGMENT	2 of 2			
SOURCE	Arabidopsis thaliana.			

ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 76620 to 83095)
AUTHORS	Parker,J.E., Coleman,M.J., Szabo,C., Frost,L.N., Schmidt,R., van der Biezen,E.A., Moores,T., Dean,C., Daniels,M.J. and Jones,J.D. The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6
TITLE	Plant Cell 9 (6), 879-894 (1997)
JOURNAL	97355983
MEDLINE	9212464
REFERENCE	2 (bases 1 to 91660)
AUTHORS	Noel,L., Moores,T.L., van Der Biezen,E.A., Parniske,M., Daniels,M.J., Parker,J.E. and Jones,J.D. Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of Arabidopsis
JOURNAL	Plant Cell 11 (11), 2099-2112 (1999)
MEDLINE	20025650
PUBMED	10559437
REFERENCE	3 (bases 76020 to 83095)
AUTHORS	Coleman,M.J., Parker,J.E., Szabo,V., van der Biezen,E.A., Daniels,M.J. and Jones,J.D.G. Direct Submission
TITLE	Submitted (10-APR-1997) Sainsbury Laboratory, John Innes Centre, Colney Lane, Norwich, Norfolk NR4 7UH, UK 4 (bases 1 to 91660)
JOURNAL	Noel,L., Moores,T., Van der Biezen,E.A., Parniske,M., Daniels,M.J., Parker,J.E. and Jones,J.D.G. Direct Submission
TITLE	Submitted (25-ANG-1999) Sainsbury Laboratory, John Innes Centre, Colney Lane, Norwich, Norfolk NR4 7UH, United Kingdom On Nov 18, 1999 this sequence version replaced gi:2109274.
COMMENT	Location/Qualifiers
FEATURES	1..91660
source	/organism="Arabidopsis thaliana" /cultivar="Landsberg erecta" /db_xref="taxon:3702" /chromosome="4" /map="C18a" 5354..9673 /gene="La-I" /pseudo join(5354..5829,5977..7057,7241..7510,7610..8071, 8281..9366,9458..9673) /gene="La-I" /note="RPP5 homolog" /pseudo /codon_start=1 15312..20023 /rpt_family="retroelement castor" /rpt_type=dispersed 15312..15431 15507..19886 /gene="encodes gag/pol" /pseudo 19904..20023 20048..22736 /gene="La-H" /pseudo /evidence=not_experimental join(20048..20440,20985..22416,22533..22736) /gene="La-H" /note="RPP5 homolog" /pseudo /codon_start=1 25823..35584 /gene="La-G" /note="interrupted by retroelement pollux1" /pseudo /evidence=not_experimental join(28283..26301,26428..27343,32765..32926,33111..33407, 33525..33899,34108..35235,35366..35584)
CDS	
gene	
LTR	
repeat_region	
gene	
LTR	
gene	
CDS	
gene	
CDS	

RESULT 13	AX032893	5895 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	AX032893				
DEFINITION	Sequence 6 from Patent WO0044918.				
ACCESSION	AX032893				
VERSION	AX032893.1	GI:10279828			
KEYWORDS					
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 5895)				
AUTHORS	Dean, C. and Gendall, A.				
TITLE	Methods and means for modification of plant characteristics using the vernalization gene vrn2				
JOURNAL	Patent: WO 0044918-A 6 03-AUG-2000;				
	DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD (GB)				
FEATURES	Location/Qualifiers				
SOURCE	1..5895				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
BASE COUNT	1742 a 1027 c 1070 g 2056 t				
ORIGIN					
Query Match	18.5%; Score 319.4; DB 6; Length 5895;				
Best Local Similarity	92.7%; Pred. No. 3.1e-60;				
Matches 370; Conservative	0; Mismatches 21; Indels 8; Gaps 3;				
QY	1308 TTCTGCGTGGAGATGTTTGTGATTAATTAATGAAGAACCATGGACCTTGTGCATCAGCC	1367			
DB					
QY	4677 TCCAGGTGTGGAGATTGTTTTGATTAACTATGGAACCAATGACCTGTCGACTCAGCC	4736			
DB					
QY	1368 ACCATCAACAACCTGCATFACCATCTCGAGAAATGCCGTATFACCTGACAGACACACACC	1427			
DB					
QY	4737 ACCATCAACAACCTGCATFACCATCTCGAGAAATGCCGTATFACCTGAG-----TCACT	4780			
DB					
QY	1428 ACCAACAACAACAACAGTGTGATGTCGCCAGTGCAGTCAACACCAACAACATFACATT	1487			
DB					
QY	4791 AACCAACAACAACAACAGTGTGATGTCGCCAGTGCAGTCAACACCAACAACATFACATT	4850			
DB					
QY	1488 GTGATATCCCATGATGATCAATTAACAAGAACAATGTGCACACAGACAAATFACAGC	1547			
DB					
QY	4851 GTGATATCCCATGATGATCAATTAACAAGAACAATGTGCACACAGACAAATFACAGC	4910			
DB					
QY	1548 AGAACAACAAGTAATTAATAGAAATCTCCGCTTTATGATACCGATTATCGGATTG	1607			
DB					
QY	4911 AGAGAC-AAGTAATTAATAGAAACACTCCGCTTTATGATACCGATTATCGGATTG	4969			
DB					
QY	1608 TAACTATCTCTCTCTTCTT-AAAAAATGTGTAGAGCAAACAATTTTATATGTTAG	1666			
DB					
QY	4970 TAACTATCTCTCTCTTCTTCTTAAAAAATGTGTAGAGCAAACAAGATTTATTTGTTAG	5029			
DB					
QY	1667 TGTAATCACTGATTACATTTTATAGTTAAAAAAAAT 1705				
DB					
QY	5030 TGTAATCACTGATTACATTTTATAGTTAAAAAAAATGAT 5068				
DB					
RESULT 14					
ATCHRIV45					
LOCUS	ATCHRIV45	199548 bp	DNA	linear	PLN 16-MAR-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 45.				
ACCESSION	AL161545				
VERSION	AL161545.2	GI:7268431			
KEYWORDS					
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 199548)				
REFERENCE	EU Arabidopsis sequencing project.				
AUTHORS					

TITLE	JOURNAL	COMMENT
Direct Submission	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemckemips.biochem.mpg.de, mayemips.biochem.mpg.de, project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk	
FEATURES	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/th1/ this fragment has an overlap with ATCHRIY44 at the 5' end and an overlap with ATCHRIY46 at the 3' end.	
source	Location/Qualifiers	
	1. 139548	
	/organism="Arabidopsis thaliana"	
	/variety="Columbia"	
	/db_xref="taxon:3702"	
	/chromosome="4"	
gene	6008. 10944	
	/gene="AT4g16810"	
gene	complement(101)(6008. 6649,6855. 7090,7222. 7258, 7337. 7536,7629. 7772,7877. 7913,8744. 8853,9020. 9108,9487. 9576,9704. 9786,9891. 9980,10178. 10229, 10883. 10944))	
	/gene="AT4g16810"	
	/note="similarity to centromere protein E, human"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="CAB80952.1"	
	/db_xref="GI:7268432"	
	/translat="MDPIKITEAKVPAKRSKATSHYLPLKRPYHSRNGOPSLIO VASDRSENDYKNDAAHLESOKLMSDBENIEVAERFTKLNSTYKOORLYADAH IWPACAEFRHLQELRSNLSLDLNETTETAPATATTHGSDGLDGEQAVV EYVVEAEFVYVVAEQAQAVVETSSSDLRMAGEDYSDTSPISBPDIINWPSYVY ESTPSSISNIEIEVDYVKNKPGVIDESIPSPEDLGKFSYVESPLDTSOG LELVYGESECVKEIOTENESPKIEGNDVCPRLLEOELVSTKVTDESSQVISEPP DLRNMSSEYKQVSPOLSDIOERGLFYSKEDLIIEESPTREGISSGIERKIKSQEIV GIGRLDSNYKENTINATDPAKVEYLDNAYSDOKMEKRSRYRLNASKDVQDESSFOE PLCLCQAEARFSPVSRINPKLSPSNDYSLHLRPIHLQESISMSNKKQSPLE QSYTEKENYKRGVSGMSNQLSSVAKETIAGKWRCPQKNKRIIVPLKRLDAWIHKV"	
exon	complement(6008. 6649)	
	/gene="AT4g16810"	
	/number=1	
Intron	complement(6650. 6854)	
	/gene="AT4g16810"	
	/number=1	
exon	complement(6855. 7090)	
	/gene="AT4g16810"	
	/number=2	
Intron	complement(7091. 7221)	
	/gene="AT4g16810"	
	/number=2	
exon	complement(7222. 7258)	
	/gene="AT4g16810"	
	/number=3	
Intron	complement(7259. 7336)	
	/gene="AT4g16810"	
	/number=3	
exon	complement(7337. 7536)	
	/gene="AT4g16810"	
	/number=4	
Intron	complement(7537. 7628)	
	/gene="AT4g16810"	
	/number=4	
exon	complement(7629. 7772)	
	/gene="AT4g16810"	
	/number=5	

```

Intron
    complement(7773..7876)
    /gene="AT4g16810"
    /number=5
exon
    complement(7877..7913)
    /gene="AT4g16810"
    /number=6
Intron
    complement(7914..8743)
    /gene="AT4g16810"
    /number=6
exon
    complement(8744..8853)
    /gene="AT4g16810"
    /number=7
Intron
    complement(8854..9019)
    /gene="AT4g16810"
    /number=7
exon
    complement(9020..9108)
    /gene="AT4g16810"
    /number=8
Intron
    complement(9109..9486)
    /gene="AT4g16810"
    /number=8
exon
    complement(9487..9576)
    /gene="AT4g16810"
    /number=9
Intron
    complement(9577..9703)
    /gene="AT4g16810"
    /number=9
exon
    complement(9704..9786)
    /gene="AT4g16810"
    /number=10
Intron
    complement(9787..9890)
    /gene="AT4g16810"
    /number=10
exon
    complement(9891..9980)
    /gene="AT4g16810"
    /number=11
Intron
    complement(9981..10177)
    /gene="AT4g16810"
    /number=11
exon
    complement(10178..10229)
    /gene="AT4g16810"
    /number=12
Intron
    complement(10230..10882)
    /gene="AT4g16810"
    /number=12
exon
    complement(10883..10944)
    /gene="AT4g16810"
    /number=13
gene
    /gene="AT4g16820"
    /number=13
CDS
    join(16254..17790,18979..19247)
    /gene="AT4g16820"
    /note="similarity to triacylglycerol lipase (EC 3.1.1.3),
    contains lipases, serine active site AA321-330
    contains EST gb:AI993282.1, T76589, AI994595.1, R30062,
    232589"
    /codon_start=1
    /product="triacylglycerol lipase like protein"
    /protein_id="CA860953.1"
    /db_xref="GI:7268433"
    /translation="MOTLPNADIPHAKRRTCTNHTSTLIPTRPLSVSPARKTNKE
    HLRLENYLTSSNSIDHTINTSROEKTKNTSTSLGGLNLRIMPOMAAADGM
    FVOAAVAFHSDDEGSPRVAIDPSFKTKSLVATSSVRLDKWIDVAIPDLRWTKO
    TPRKVEGSLTYTDOHAPSLAESLVEISRLVELVAGELISVYGHSGAIAL
    LAADDAIERPAPVAVAFSGFRVGNNEFADRLDSKGVKLVNVSQDSVATKPGI
    HDLEAVLHVDGFIASNCBPRANAKRSKLKLDKORSNVKLYTGKSLRLNRTFHDG
    DVLPRDTGGDNDPIKYYKLFRIAVGLSLITTVFLPIPNHFFYFHFLLIFLSLFL
    SLNQGVLSLTHFVSITLFSIDLILISMAWSSH"
    16254..17790

```

```

Intron
    /gene="AT4g16820"
    /number=1
    17791..18978
    /gene="AT4g16820"
    /number=1
exon
    18979..19247
    /gene="AT4g16820"
    /number=2
gene
    19125..20999
    /gene="AT4g16830"
    /number=2
CDS
    19101,19312..19483,19652..19958,20041..20106,20210..20336,
    20437,20592,20679..20729,20811..20999)
    /gene="AT4g16830"
    /note="cognate cDNA CDNA4KMY is matching at position
    similarity to nuclear antigen - Bos taurus
    contains EST gb:AI994595.1, AI993282.1, R30062, 742546,
    232590, 232589, T76589, AA395095"
    /codon_start=1
    /product="nuclear antigen homolog"
    /protein_id="CA860954.1"
    /db_xref="GI:7268434"
    /translation="MOTLPNADIPHAKRRTCTNHTSTLIPTRPLSVSPARKTNKE
    HLRLENYLTSSNSIDHTINTSROEKTKNTSTSLGGLNLRIMPOMAAADGM
    FVOAAVAFHSDDEGSPRVAIDPSFKTKSLVATSSVRLDKWIDVAIPDLRWTKO
    TPRKVEGSLTYTDOHAPSLAESLVEISRLVELVAGELISVYGHSGAIAL
    LAADDAIERPAPVAVAFSGFRVGNNEFADRLDSKGVKLVNVSQDSVATKPGI
    HDLEAVLHVDGFIASNCBPRANAKRSKLKLDKORSNVKLYTGKSLRLNRTFHDG
    DVLPRDTGGDNDPIKYYKLFRIAVGLSLITTVFLPIPNHFFYFHFLLIFLSLFL
    SLNQGVLSLTHFVSITLFSIDLILISMAWSSH"
    19312..19483
    /gene="AT4g16830"
    /number=1
    19484..19651
    /gene="AT4g16830"
    /number=1
    19652..19958
    /gene="AT4g16830"
    /number=2
    19959..20040
    /gene="AT4g16830"
    /number=2
    20041..20106
exon
    1308 TTCTGCTGTGGAGATGTTTGTGATTAACATGAGAACCATGACTGTCAGCTGAGCC 1367
    28171 TCACAGTGTGGAGATGTTTGTGATTAACATGAGAACCATGACTGTCAGCTGAGCC 28230
QY 1368 ACCATCAACCACTCAATACCATCTCGAGAAATGCCGTAAATGCGACAGACACACGACC 1427
    28231 ACCATCAACCACTCAATACCATCTCGAGAAATGCCGTAAATGCGACAGACACACGACC 28284
QY 1428 ACCAACAACAACAACAGCTGAGATCGTCCAGAGACTCAACAACCAACAACAATTAACAT 1487
    28285 AACACAACAACAACAACAGCTGAGATCGTCCAGAGACTCAACAACCAACAACAATTAACAT 28344
QY 1488 GTGATCATCCCAATGACATTAACAACAACAACAACAATGTTGACACAAGGACATTAACGC 1547
    28345 GTGATCATCCCAATGACATTAACAACAACAACAACAATGTTGACACAAGGACATTAACGC 28404
QY 1548 AGAGACAACAACAATTAAGAAATATCCGCGCTTTATGATACCATTTATGATGATG 1607
    28405 AGAGAC-AACTAATTTAAATAGAAACACACGCGTTTATGATGATGATGATGATG 28463
QY 1608 TAACTTAATCTCTTCTTCTT-AAAAAATGTTAGAGAGAAACAATTTTATATGATG 1666
    28464 TAACTTAATCTCTTCTTCTTCTTAAAAAATGTTAGAGAGAAACAATTTTATATGATG 28523
QY 1667 TGTATTCACATGATTAATTTTATGTTAAAAAATAAT 1705
    ||||||||||||||||||||||||||||||||||||

```

Query Match 18.5% Score 319.4; DB 8; Length 199548;
 Best Local Similarity 92.7%; Pred. No. 3,1e-60;
 Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;


```
repeat_region /number=1
               /note="23bp T tandem repeat"
               /number=16170
misc_feature 17691..19115
              /note="CDNA1KM"
              /number=17793..19171
gene /gene="d14425c"
      /complement(join(17793..18046,18147..18250,18348..18399,
18481..18571,18886..18960,19058..19171))
      /gene="d14425c"
      /complement(join(17793..18046,18147..18250,18348..18399,
18481..18571,18886..18960,19058..19171))
      /note="strong similarity to enoyl-CoA hydratase (EC
4.2.1.17), AU-specific RNA-binding protein, human"
      /codon_start=1
      /product="enoyl-CoA hydratase"
      /protein_id="CAB10453.1"
      /db_xref="GI:2245034"
      /db_xref="SPTREMBL:O23520"
      /translation="MTKSLONAFESITHONSAHYVMIRSLVPGVCAGADLKERRTMS
PEVHTIVNSLRITWESFTEALSTPTTAAIEGALIGLEPMALACDLRIGEMAVVAGLP
ETGALITPGAGGTGRLSRVGRSVSKELITGKIDAIENAKGLVNTICVTGPAHEK
AIEHAQDINERGLAIEMAKAIDEGITETNMASGLEVEEMCTOKLINTODRLEGIAAF
AEKRRLPLYTGN"
exon /complement(17793..18046)
      /gene="d14425c"
      /number=1
      /complement(18047..18146)
      /gene="d14425c"
      /number=1
      /complement(18147..18250)
      /gene="d14425c"
      /number=2
      /complement(18251..18347)
      /gene="d14425c"
      /number=2
      /complement(18348..18399)
      /gene="d14425c"
      /number=3
      /complement(18400..18480)
      /gene="d14425c"
      /number=3
      /complement(18481..18571)
      /gene="d14425c"
      /number=4
      /complement(18572..18885)
      /gene="d14425c"
      /number=4
      /complement(18886..18960)
      /gene="d14425c"
      /number=5
      /complement(18961..19057)
      /gene="d14425c"
      /number=5
      /complement(19058..19171)
      /gene="d14425c"
      /number=6
      /number=25115
gene
```

Query Match 18.5%; Score 319.4; DB 8; Length 201471;
Best Local Similarity 92.7%; Pred. No. 3,1e-60;
Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;

```
QY 1308 TTCGCTGTGGAGTGGTTTGTATTAACATGAGCAACATGAGCTTGTCGACTCAGCC 1367
DB 42342 TCAGAGTGTGAGATTTTGTATTAACATGAGCAACATGAGCTTGTCGACTCAGCC 42401
QY 1368 ACCATCAACACTGCATACATCTCTGAGAAATGGCGTAATAGCTCAGACACACAGCC 1427
DB 42402 ACCATCAACACTGCATACATCTCTGAGAAATGGCGTAATAGCTCAGACACACAGCC 1427
QY 1428 ACCAACAACACACATGATGATGTCAGTGCAGTCAAAACACACACATTAACATT 1487
```

```
DB 42456 AACACACACACACACACATGTGATCATCCAGTCAACACACACACATTAACATT 42515
QY 1488 GTGATCATCCCATGACATTAACACACACACATGTTGACACACACACACATTAACAGC 1547
DB 42516 GTGATCATCCCATGACATTAACACACACACATGTTGACACACACACACATTAACAGC 42575
QY 1548 AGACACACACACACACATTAACATGAAATTCGCGCTTTATGATACGATTTATCGATTG 1607
DB 42576 AGAGAC-AGATTAATTAATGAGAAACACCTCGCTTTAGATGATACGATTTATCGATTG 42634
QY 1608 TAACCTATTCCTCTTCTT-AAAAATGTTTGGAGACACAAATTTTATATGATTG 1666
DB 42635 TAACCTATTCCTCTTCTTAAAAAATGTTTGGAGACACAAATTTTATATGATTG 42694
QY 1667 TGTATTCACATGATTACATTTTATGTTAAAAAATAAT 1705
DB 42695 TGTATTCACATGATTACATTTTATGTTAAAAAATGAT 42733
```

Search completed: June 19, 2003, 21:25:47
Job time : 4500 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 17:43:25 ; Search time 422 Seconds

9189.434 Million cell updates/sec

Title: US-09-890-220-1

Sequence: 1 caagcttctcaatttgct.....aatgattctgctataact 1722

Scoring table: IDENTITY_NUC

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Maximum DB seq length: 20000000000

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002: *

1:	/SID52/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2:	/SID52/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SID52/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SID52/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SID52/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SID52/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SID52/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SID52/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SID52/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SID52/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SID52/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SID52/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SID52/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SID52/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SID52/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SID52/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SID52/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SID52/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SID52/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SID52/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SID52/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24:	/SID52/gcgdata/genseq/genseqn-emb1/NA2001C.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1722	100.0	1722	21	AAA65670	Nucleotide sequen
2	1722	100.0	1722	21	AAA47751	VRN2 nucleic acid.
3	1569.6	91.1	1715	21	AAA67571	Nucleotide sequen
4	1569.6	91.1	1715	21	AAA47753	VRN2 nucleic acid.
5	1537.6	89.3	1737	21	AAA57581	aberrantly splice
6	979.8	56.9	1497	21	AACT5155	Arabidopsis thalia
7	412.2	23.9	6338	21	AAA47752	VRN2 nucleic acid.
8	319.4	18.5	5895	21	AAA47754	VRN2 nucleic acid.
9	314.2	18.2	319	21	AACT5152	Arabidopsis thalia

C	10	306.4	17.8	640	21	AAAC44147	Arbidopsis thailaia
	11	291.4	16.9	2280	21	AAAC36947	CDNA encoding a MP
	12	252.4	14.7	2248	21	AAAZ36953	CDNA encoding an O
	13	230.8	13.4	558	21	AAAF47758	At Hvp 2245035 (AT
	14	170.8	9.9	525	21	AAAC54919	Arbidopsis thailaia
	15	163.8	9.5	325	21	AAAA47757	At161743 Expressed
	16	126.6	7.4	463	21	AAAA47756	Rice expressed seq
	17	71	4.1	8045	21	AAA651171	Neurospora crassa
C	18	69.6	4.0	567	21	AAA29550	HIV codon altered
	19	69	4.0	198	17	AAAT1204	DNA-spanner oligon
	20	69	4.0	198	17	AAAT17205	DNA-spanner oligon
	21	65.8	3.8	198	17	AAAT17202	DNA-spanner oligon
	22	65.8	3.8	198	17	AAAT17203	DNA-spanner oligon
	23	63.2	3.7	1216	21	AAAZ36962	DNA fragment encod
	24	62.2	3.6	5580	21	AAAZ36948	DNA encoding a MPC
	25	62	3.6	5511	21	AAA61847	Cryptosporidium pa
	26	62	3.6	5511	21	ABT04776	C parvum GP900 gen
	27	62	3.6	7334	24	AAA61846	Cryptosporidium pa
	28	62	3.6	7334	24	ABT04775	C parvum GP900 gen
	29	60.4	3.5	5163	19	AAV20700	Cryptosporidium pa
	30	60.4	3.5	5163	21	AAA61849	ORF encoding a por
	31	60.4	3.5	5163	21	ABT04778	C parvum GP900 gen
	32	60.4	3.5	5318	19	AAAT20701	Cryptosporidium pa
	33	60.4	3.5	5318	21	AAA61848	DNA encoding a por
	34	60.4	3.5	5318	24	ABT04777	C parvum GP900 gen
C	35	59	3.4	6120	22	AAAS46788	Tumour suppressor
	36	56.8	3.3	2336	22	ABL25662	Drosophila melanog
	37	55.8	3.2	3300	21	AAZ55692	DNA encoding yeast
	38	54.6	3.2	4829	23	ABLI13399	Drosophila melanog
C	39	54.6	3.2	39746	23	ABLI13398	Drosophila melanog
C	40	54.2	3.1	8524	22	AAAS46588	Tumour suppressor
C	41	54.2	3.1	8524	24	ABL33841	Human immune syste
	42	54	3.1	144	12	AAO12515	CSP-2 peptide from
	43	53.4	3.1	1590	24	AAQ93524	Nicotiana glauca ar
	44	53.2	3.1	2892	24	ABN79885	Fungal ZBC gene se
C	45	52.2	3.0	8662	24	ABL34637	Human metastasis a

ALIGNMENTS

PT	Arabidopsis, encoding a polypeptide capable of specifically altering
PT	the flowering time of a plant
PT	
DR	
DR	WPI; 2000-532899/48.
PI	Johanson U, West J, Dean C;
PI	
XX	
XX	
XX	(PLAN-) PLANT BIOSCIENCE LTD.
XX	
XX	
XX	05-FEB-1999; 99GB-0002660.
PR	
XX	
XX	25-JAN-2000; 2000WO-GB00197.
PF	
XX	
XX	
PD	10-AUG-2000.
PN	
PN	WO200046358-A2.
XX	
XX	
OS	Arabidopsis sp.
XX	
XX	
KW	flower initiation; stem elongation; flower production; VRN2; ss.
KW	
XX	
XX	
DE	Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
DE	
DT	
DT	04-DEC-2000 (first entry)
AC	
AC	AAA63670;
ID	AAA63670 standard; cDNA; 1722 BP.
RESULT 1	
AAA63670	

XX Example 5; Page 53-54; 73pp; English.

The present sequence represents a VRN2 sequence from variety Landsberg erecta. The specification describes a sequence which encodes a polypeptide capable of specifically altering the flowering time of a plant. The polypeptide is encoded by the FRI (one locus-FRI₁) locus of Arabidopsis. The FRI polynucleotide is used to transform plants, so that the flowering time of a plant is altered. This is used, for example for plants in which the leaves or tubers are a commercial product, where it is desirable to avoid 'blotting' (initiation of flowers and stem elongation) at too early a stage. Conversely, it may be desirable to alter flowering under certain circumstances e.g. to vary flower production across the seasons.

SQ Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

Query Match	100.0%	Score 1722;	DB 21;	Length 1722;
-------------	--------	-------------	--------	--------------

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Indels 0; Gaps 0;

QY	1	CAAGCTCTTCGAATTTGGTCTGCTGCTCTTACGACCGAAATCGGCTTTTGGACGCTT	60
Dp	1	CAGCTCTTTCGAATTTGGTCTGCTCTCTTACGACCGAAATCGGCTTTTGGACGCTT	60
QY	61	CAGGCTCAATCCAAAGACATCTATATATAGCATATTCGAGAAAGCGGCTCTAATGTT	120
Dp	61	CAGGCTCAATCCAAAGACATCTATATATAGCATATTCGAGAAAGCGGCTCTAATGTT	120
QY	121	GCATGAGTTTATCGCTATAGCTAGGAGAAATTTCAATTTAGGGAGCGCTCGAGATTG	180
Dp	121	GCATTAGTTTATCGCTATAGCTAGGAGAAATTTCAATTTAGGGAGCGCTCGAGATTG	180
QY	181	CACCTACTTCATATACGGCTCTGCTGACGTTGTTGAGCTAATTTAACAAGATGTAGGC	240
Dp	181	CACCTACTTCATATATACGGCTCTGCTGACGTTGTTGAGCTAATTTAACAAGATGTAGGC	240
QY	241	AGAAATGTGCGCGAAATCTGCACCGGAGAAAGATTTCAACGATGAGAAATCTGTGA	300
Dp	241	AGAAATGTGCGCGAAATCTGCACCGGAGAAAGATTTCAACGATGAGAAATCTGTGA	300
QY	301	TATATTTGAACCTGTGCTGACTATATTAACATCTTCACTGCTGCTCTAGGAAACCAT	360
Dp	301	TATATTTGAACCTGTGCTGACTATATTAACATCTTCACTGCTGCTCTAGGAAACCAT	360
QY	361	CGTTCTTCCAGATAGCTGACATCTGAACCTAAATTTGAGCAAAAGCGCAAAAGAAATCTCAAGT	420
Dp	361	CGTTCTTCCAGATAGCTGACATCTGAACCTAAATTTGAGCAAAAGCGCAAAAGAAATCTCAAGT	420
QY	421	CTACTCTGGATGCTAGTTTTCACATTAAGAGTTGTAATTAACATTTACGAAAACTGAAG	480
Dp	421	CTACTCTGGATGCTAGTTTTCACATTAAGAGTTGTAATTAACATTTACGAAAACTGAAG	480
QY	481	TTAGGAGAGATTTGTTCTTGCCATTTTCTCTATGCTATGCTATGTGGTACCTTCAGGGCTGC	540
Dp	481	TTAGGAGAGATTTGTTCTTGCCATTTTCTCTATGCTATGCTATGTGGTACCTTCAGGGCTGC	540
QY	541	AATTTCAATTTGAATTCATCTCATGATGATTTATTTGAAATTTGAGTTCAAGCTTTTCSAAGAT	600
Dp	541	AATTTCAATTTGAATTCATCTCATGATGATTTATTTGAAATTTGAGTTCAAGCTTTTCSAAGAT	600
QY	601	ACCAGACATTTAATGTTCTCTGTAATAAACTAATTCCTTCATATTTGAGAAAGAAAGAGTG	660
Dp	601	ACCAGACATTTAATGTTCTCTGTAATAAACTAATTCCTTCATATTTGAGAAAGAAAGAGTG	660
QY	661	ATGAGCATTAATTTAGCGCTTCTCTCTGCTCTGCAAACTCGTAAGCGGAGACAAAGAG	720
Dp	661	ATGAGCATTAATTTAGCGCTTCTCTCTGCTCTGCAAACTCGTAAGCGGAGACAAAGAG	720
QY	721	GTGGCGAAATTAACACCGAGAGCTTAAAGTATGCTTTTACGGTGGATTCACCCAGTT	780
Dp	721	GTGGCGAAATTAACACCGAGAGCTTAAAGTATGCTTTTACGGTGGATTCACCCAGTT	780

OY	78	TACCTAATGCGACAGAAAATGGAAATCAACCCTACTTAATGATGGAACCCGTGGTTAGAT	840
Db	781	TACTAATGCGACAGAAAATGGAAATCAACCCTACTTAATGATGGAACCCGTGGTTAGAT	840
OY	841	ATCCGAGGCACAACAGCTTGCTGGACAAATTGGATGACACACCAATTCACACAGCA	900
Db	841	ATCCGAGGCACAACAGCTTGCTGGACAAATTGGATGACACACCAATTCACACAGCA	900
OY	901	TACCCACACTCTTCTCTGSGACCGTGGTCTAAAGTATATTTGACAAACCGAAGCTGTGGTC	960
Db	901	TACCCACACTCTTCTCTGSGACCGTGGTCTAAAGTATATTTGACAAACCGAAGCTGTGGTC	960
OY	961	CTGCTACTAAGAACAAAGAAGTTATCTGCTGAGCGATCAGAGCTTAGAGCCACTTATCTTC	1020
Db	961	CTGCTACTAAGAACAAAGAAGTTATCTGCTGAGCGATCAGAGCTTAGAGCCACTTATCTTC	1020
OY	1021	TTTCAGAAAGGCCAATTCATCATCTCTCACAGATCCAGCCAATGGGGCTTTGAGCAAGTAA	1080
Db	1021	TTTCAGAAAGGCCAATTCATCATCTCTCACAGATCCAGCCAATGGGGCTTTGAGCAAGTAA	1080
OY	1081	TGTCTACACGGGATAGCAGAGTGAAGTCGATGCGATGTGGAGATTTTGAAGATCGCC	1140
Db	1081	TGTCTACACGGGATAGCAGAGTGAAGTCGATGCGATGTGGAGATTTTGAAGATCGCC	1140
OY	1141	AGATGGTTATGACTTTGGATGGTGAATAAGATGAAAGCAATTCAATGATCTTTGGA	1200
Db	1141	AGATGGTTATGACTTTGGATGGTGAATAAGATGAAAGCAATTCAATGATCTTTGGA	1200
OY	1201	ACACGTTTTAAGAAAAAAGGGTTATAGAGATGGTCAATCTTTGGGATGTGAG	1260
Db	1201	ACACGTTTTAAGAAAAAAGGGTTATAGAGATGGTCAATCTTTGGGATGTGAG	1260
OY	1261	CATTTTCAAAGATTTTACAGAAAAGATGGTCAACCGTACTACTATCATCTTTCTGTGTGGA	1320
Db	1261	CATTTTCAAAGATTTTACAGAAAAGATGGTCAACCGTACTACTATCATCTTTCTGTGTGGA	1320
OY	1321	GATTGTTTTGATTTAACTATGAGAACATGGACTTTGTGCATCGACACCATCAACAACACT	1380
Db	1321	GATTGTTTTGATTTAACTATGAGAACATGGACTTTGTGCATCGACACCATCAACAACACT	1380
OY	1381	GCATATCCATCCTCAGAAATGGCGTAAATAGCTAGACACACACACACCAACAACA	1440
Db	1381	GCATATCCATCCTCAGAAATGGCGTAAATAGCTAGACACACACACACCAACAACA	1440
OY	1441	ACAGTGGATGTCGCCACGATCAACAAACCAACCAATTAACATTTGGATTCATCCCA	1500
Db	1441	ACAGTGGATGTCGCCACGATCAACAAACCAACCAATTAACATTTGGATTCATCCCA	1500
OY	1501	ATGACATAAACAACAAAGAAATGTTGACAAACAAAGACATTAACACAGACAGACAAAGTAA	1560
Db	1501	ATGACATAAACAACAAAGAAATGTTGACAAACAAAGACATTAACACAGACAGACAAAGTAA	1560
OY	1561	TTAATAAGAGAAATATCTCCGGCTTTTATGATACGATTTATCGATTTTACTTATTTCTTC	1620
Db	1561	TTAATAAGAGAAATATCTCCGGCTTTTATGATACGATTTATCGATTTTACTTATTTCTTC	1620
OY	1621	TTTTCTTAAAAAAATTTGTTAGAGCAACAATTTTTTATATGTTAGTGTATTCAACTGAT	1680
Db	1621	TTTTCTTAAAAAAATTTGTTAGAGCAACAATTTTTTATATGTTAGTGTATTCAACTGAT	1680
OY	1681	TACATTTTATGTTAAAAAAAATGATATCTGCTTATTACT 1722	
Db	1681	TACATTTTATGTTAAAAAAAATGATATCTGCTTATTACT 1722	
<hr/>			
RESULT 2	AAA47751	standard; cDNA: 1722 BP.	
ID	AAA47751		
XX	AAA47751:		
AC			
DT	16-NOV-2000	(first entry)	

RESULT 2

ID AAA47751 standard; cDNA; 1722 BP.

AC AAAA47751;

DT 16-NOV-2000 (first entry)

XX

DE VRN2 nucleic acid.

XX Vernalization gene; VRN2; plant characteristic; flowering time;

KW leaf size; leaf shape; shade avoidance response; reproduction;

KW breeding; pollination; cultivation; ss.

XX Arabidopsis thaliana var 'Landsberg erecta'.

OS

XX

Key location/Qualifiers

FT CDS 231..1568

FT /*tag= a

FT /product= "VRN2 polypeptide"

FT replace(1159,A)

FT mutation

FT /*tag= b

FT /note= "found in mutant vrn2-1; alters a Trp codon

FT to a STOP codon leading to a truncation of

FT the protein"

XX WO200044918-A1.

XX PD 03-AUG-2000.

XX 28-JAN-2000; 2000WO-GB00248.

XX PR 28-JAN-1999; 99GB-0001927.

XX (PLAN-) PLANT BIOSCIENCE LTD.

XX PA

XX Dean C. Gendall A;

XX PI WPI; 2000-499333/44.

XX DR P-PSDB; AAB00060.

XX

XX Isolated vernalization gene VRN2 is used to produce transgenic plants

PT with altered vernalization response, flowering time, leaf size and/or

PT shape or shade avoidance response for maximized reproductive success

XX

PS Claim 4; Fig 6; 105pp; English.

XX

XX Isolated nucleic acid sequences obtained from the VRN2 locus of a

CC plant encode polypeptides which are capable of affecting one or more

CC vernalization responses such as, flowering time, leaf size and/or

CC shape or the shade avoidance response of a plant into which the

CC nucleic acid is introduced. Introducing such sequences into plants

CC to alter these characteristics maximises the reproductive success of

CC the plant.

XX

XX Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

SO

Query Match 100.0%; Score 1722; DB 21; Length 1722;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCTTCTTCGAATTTGCTGCTCTCTTACACAGCCAAATCGTGTGTTTCGAGCTTT 60

DB 1 CAAGCTTCTTCGAATTTGCTGCTCTCTTACACAGCCAAATCGTGTGTTTCGAGCTTT 60

OY 61 CAGGCTTCATCCAAAGACATCTATATAGCATATTTGAGAAGAGCGGCTCTAATTTT 120

DB 61 CAGGCTTCATCCAAAGACATCTATATAGCATATTTGAGAAGAGCGGCTCTAATTTT 120

OY 121 GCATTGAGTTATGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGAGTTTG 180

DB 121 GCATTGAGTTATGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGAGTTTG 180

OY 121 GCATTGAGTTATGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGAGTTTG 180

DB 121 GCATTGAGTTATGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGAGTTTG 180

OY 181 CACTTAATCTCAATTCGAGCTCTGACGTTGTTGAGTGAATTTGAACAAGAAATGTGTAGC 240

DB 181 CACTTAATCTCAATTCGAGCTCTGACGTTGTTGAGTGAATTTGAACAAGAAATGTGTAGC 240

OY 241 AGAATTTGCGCGGAATCTCAGCGGAGGAAGTATTTCAACTGATAGATCTCTTGA 300

DB 241 AGAATTTGCGCGGAATCTCAGCGGAGGAAGTATTTCAACTGATAGATCTCTTGA 300

OY 301 TATATTGTAACCTGTGACATATATTAACATTTTCACCTTGGCTCTAGGCAACCCAT 360

DB 301 TATATTGTAACCTGTGACATATATTAACATTTTCACCTTGGCTCTAGGCAACCCAT 360

OY 361 CGTTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGGCAAGAAAGTCAAGAT 420

DB 361 CGTTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGGCAAGAAAGTCAAGAT 420

OY 421 CTACTGGATGCTGATTTTCAACTATTAAGATTTGAATTAACATTTACAGAAACTGANG 480

DB 421 CTACTGGATGCTGATTTTCAACTATTAAGATTTGAATTAACATTTACAGAAACTGANG 480

OY 481 TTAGGAGGATTTGCTGCTGATTTTCCATTTTCTGATGCTGATGCTGATGCTGATGCTG 540

DB 481 TTAGGAGGATTTGCTGCTGATTTTCCATTTTCTGATGCTGATGCTGATGCTGATGCTG 540

OY 541 AATTTCAATTTGATTCATCTCATGATTTTATTTGAATTTGATTTGATTTGATTTGATTTG 600

DB 541 AATTTCAATTTGATTCATCTCATGATTTTATTTGAATTTGATTTGATTTGATTTGATTTG 600

OY 601 ACCAGACGTTAATGTTTCTGTAATTAACCTTAATTCCTCATATTTGAGGAAGAGAGTG 660

DB 601 ACCAGACGTTAATGTTTCTGTAATTAACCTTAATTCCTCATATTTGAGGAAGAGAGTG 660

OY 661 ATGACGATTAATTTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

DB 661 ATGACGATTAATTTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

OY 721 GTGGCAGAAATTAACACCGAGAGACTTAAAGTATGCTTTTACCGTTGATTCACCAAGT 780

DB 721 GTGGCAGAAATTAACACCGAGAGACTTAAAGTATGCTTTTACCGTTGATTCACCAAGT 780

OY 781 TAACATAAGCAGCAAGAAATGGAATCACCTTACTTAATGATGAGAAACCGTGTAGAT 840

DB 781 TAACATAAGCAGCAAGAAATGGAATCACCTTACTTAATGATGAGAAACCGTGTAGAT 840

OY 841 ATCCGAGCAACAGAGCTTCTGAGCAATTTGAGATGACCAACATTCACCAAGCA 900

DB 841 ATCCGAGCAACAGAGCTTCTGAGCAATTTGAGATGACCAACATTCACCAAGCA 900

OY 901 TAGCCACTCTTCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

DB 901 TAGCCACTCTTCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

OY 961 CTGCTACTAAGCAAGAAAGTTATCTGCTGAGCCTGAGAGCTGAGAGCCACTACTTC 1020

DB 961 CTGCTACTAAGCAAGAAAGTTATCTGCTGAGCCTGAGAGCTGAGAGCCACTACTTC 1020

OY 1021 TTCAGAAACGCCAATTTATCATTTCTACAGAGTCCAGCAATGGCGCTTGAAGCAATGA 1080

DB 1021 TTCAGAAACGCCAATTTATCATTTCTACAGAGTCCAGCAATGGCGCTTGAAGCAATGA 1080

OY 1081 TGCTGACCGGGATGAGCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1140

DB 1081 TGCTGACCGGGATGAGCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1140

OY 1141 AGATGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200

DB 1141 AGATGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200

OY 1201 ACTGCTTGTGAAGAAACAAAGGCTTATAGCAGATGATGATGATGATGATGATGATGATG 1260

DB 1201 ACTGCTTGTGAAGAAACAAAGGCTTATAGCAGATGATGATGATGATGATGATGATGATG 1260

OY 1261 CATTTTCAAGATTTTACGAGAAAGATGACCGTTACTCATCTCTTCTGAGTGTGGA 1320

DB 1261 CATTTTCAAGATTTTACGAGAAAGATGACCGTTACTCATCTCTTCTGAGTGTGGA 1320

OY 1321 GATTTGTTTATTAATATGAGAACCATGATGATGATGATGATGATGATGATGATGATGAT 1380

DB 1321 GATTTGTTTATTAATATGAGAACCATGATGATGATGATGATGATGATGATGATGATGAT 1380

OY 1381 GCAATATCATCTCTGAGAAATTTGCGTAAATAGCTGAGAACCAACCAACCAACCAACCA 1440

D _b	1381	GCATATACCATTCCTCGAGAAATTGGCGCTAATAGCTCGACACACCACCACCAACCAACAA	1440
Q _Y	1441	ACAGTGTGATTCGTCCAGTCAGTACTCAACACCAACAACAAATTAACATTGTGGATCATCCCA	1500
D _b	1441	ACAGTGTGATTCGTCCAGTCAGTACTCAACACCAACAACAAATTAACATTGTGGATCATCCCA	1500
Q _Y	1501	ATGACATTAACCAACAAAGAACAAATGTTGACAAACAAGACAATTAACAGACAGACACAAATTA	1560
D _b	1501	ATGACATTAACCAACAAAGAACAAATGTTGACAAACAAGACAATTAACAGACAGACACAAATTA	1560
Q _Y	1561	TTAAATAGGAAATCTCCGGCTTTTATGATACCGATTATTCGATTTTAACTTATTTCTTC	1620
D _b	1561	TTAAATAGGAAATCTCCGGCTTTTATGATACCGATTATTCGATTTTAACTTATTTCTTC	1620
Q _Y	1621	TTTTCTTAAAAAATTTGTTTAGAGCAACAATTTTTTATATGTTAGTGTATTCACATGAT	1680
D _b	1621	TTTTCTTAAAAAATTTGTTTAGAGCAACAATTTTTTATATGTTAGTGTATTCACATGAT	1680
Q _Y	1681	TACATTTTATGTTAAAAAATAAATGAGATTCTGCTTTAACT	1722
D _b	1681	TACATTTTATGTTAAAAAATAAATGAGATTCTGCTTTAACT	1722

RESULT 3

ID	AAA63671	standard; cDNA; 1715 BP.
1	1	1

AC AAA63671;

DT 04-DEC-2000 (first entry)
XX

Nucleotide sequence of VRN2 sequence of Arabidopsis sp.

KW H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;
 KW flower initiation; stem elongation; flower production; VRN2; ss

OS Arabidopsis sp.

PN WO200046358-
XX

PD 10-AUG-2000.
XY

PF 25-JAN-2000; 2000WO-GB00197
 YY

PR 05-FEB-1999; 99GB-0002660
YY

PA (PLAN-) PLANT BIOSCIENCE LTD.
XY

PI Johanson U, West J, Dean C;
XX

DR WPI; 2000-532899/48.
XX

PT New nucleic acid derived from the FRI locus of a plant, e.g. Arabidopsis, encoding a polypeptide capable of crosslinking PT

pt the flowering time of a plant -
xx

PS Example 5; Page 54-55; 73pp; English.
XX

CC The present sequence represents a VRN2 sequence from variety Columbia
CC The specification describes a sequence which encodes a polypeptide.
CC

The polypeptide is encoded by the ERT (one locus-ERTIDA) locus

of Arabidopsis. The FRI polynucleotide is used to transform plants, so that the flowering time of a plant is altered. This is used, for example, CC CC

for plants in which the leaves or tubers are a commercial product, where it is desirable to avoid 'blotting' /initiation of flowers and stem

CC elongation) at too early a stage. Conversely, it may be desirable to
CC alter flowering under certain circumstances e.g. to vary flower

cc production across the seasons.
xx

35. sequence 1/13 BF; 336 A; 326 C; 353 G; 500 T; 0 other;

[illegible]

Db		1021	TCTTCAGAAAGCCCAATTCTATCTCATTTCTCACAGAGTCCAGCATTGGCGCTTGAGCAACT	1080		
OY		1079	AATGTCGTACCGGATAGCGAGATGAAGTGCATGACGATGTTGCAATTTTGAAGATCG	1138		
Db		1081	AATGTCGTATCGGATAGCGAGATGAAGTGCATGACGATGTTGCAATTTTGAAGATCG	1140		
OY		1139	CCCAATGCGTTGATGACTTTGGATGTGAATTAAGATGAAGAAGCAATTCATGATCTTTGG	1198		
Db		1141	CCAATAGCTTGTATGACTTTGTGGATGTGAATAAAGATGAAGAAGCAATTCATGATCTTTGG	1200		
OY		1199	GAACTCGTTTGTAGAAAACAACAAGCGTTATAGCAGATGAGTCATATCTCTTGGCGATGGA	1258		
Db		1201	GAACTCGTTTGTAGAAAACAACAAGCGTTATAGCAGATGAGTCATATCTCTTGGCGATGGA	1260		
OY		1259	AGCATTTTTCAGATTTTACGAGAAAGATTGCCACCGTTACTCATCACTCTTTGGTGTGG	1318		
Db		1261	AGTATTTTTCAGATTTTACGAGAAAGATTGCCACCGTTACTCATCACTCTTTGGTGTGG	1320		
OY		1319	GAGATGTGTTTTGATTTAAACATATGGAACCATATGACTGTGCAGTCAGCACCATCAACAA	1378		
Db		1321	GAGATGTGTTTTGATTTAAACATATGGAACCATATGACTGTGCAGTCAGCACCATCAACAA	1380		
OY		1379	CTGCAATATCCATCTCTCGAGATTTCCGTAATAGTCTCAGACACCACACCAACAAACAA	1438		
Db		1381	CTGCAATATCCATCTCTCGAGATTTCCGTAATAGTCTCAGACACCACACCAACAAACAA	1434		
OY		1439	CACAGATGTGATGTGCCAGTAGCTCAACACACCACAAACAATTAACATTTGGATCATCC	1498		
Db		1435	CACAGATGTGATGTATCCCACTGCTCAACACACCACAAACAATTAACATTTGGATCATCC	1494		
OY		1499	CAATGACATTAACAACAAGAACATGTTGACACACAGACAAATTAACAGACAGACAAAGT	1558		
Db		1495	GAAAGACATTAAAAAACAAGAACATGTTGACACACAGACAAATTAACAGACAGACAAAGT	1553		
OY		1559	AATTAATATAGAAATATCCGGCTTTTATGATATACCATTTATTCGATTTGTAATTAATCT	1618		
Db		1554	AATTAATATAGAAATATCCGGCTTTTATGATATACCATTTATTCGATTTGTAATTAATCT	1613		
OY		1619	TCTTTCTTT-AAAAAATGTTTAGAGGACAAACAATTTTATATGTTAGTGTATCAACT	1677		
Db		1614	TCTTTCTTTAAAAAATGTTTAGAGGACAAACAAGATTTTATTTGTTAGTGTATCAACT	1673		
OY		1678	GATTACATTTTGTAGTAAAAAAAAT 1705			
Db		1674	GATTACATTTTGTAGTAAAAAAAATGAT 1701			
RESULT 4						
ID	AAA47753	standard; cDNA; 1715 BP.				
XX	AAA47753:					
DE	16-NOV-2000	(first entry)				
XX	VRN2 nucleic acid.					
KW	Vernalization gene; VRN2; plant characteristic; flowering time;					
KW	leaf size; leaf shape; shade avoidance response; reproduction;					
KW	breeding; pollination; cultivation; ss.					
OS	Arabidopsis thaliana var 'Columbia'.					
XX	WO200044918-A1.					
PN	03-AUG-2000.					
PD	28-JAN-2000; 2000WO-GB00248.					
FR	28-JAN-1999; 99GB-0001927.					
PA	(PLAN-) PLANT BIOSCIENCE LTD.					

xx	Dean C, Gendall A;
PI	
xx	
DR	WPI: 2000-499333/4.
DR	P-PSDB: AAB00061.
xx	
PT	Isolated vernalization gene VRN2 is used to produce transgenic plants
PT	with altered vernalization response, flowering time, leaf size and/or
XX	shape or shade avoidance response for maximized reproductive success
xx	
PS	Claim 5; Page 71; 105pp; English.
xx	
CC	Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC	plant encode polypeptides which are capable of affecting one or more
CC	vernalization responses such as, flowering time, leaf size and/or
CC	shape or the shade avoidance response of a plant into which the
CC	nucleic acid is introduced. Introducing such sequences into plants
CC	to alter these characteristics maximises the reproductive success of
CC	the plant.
xx	
SQ	Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other:
	Query Match 91.1%; Score 1569, 6; DB 21; Length 1715;
	Best Local Similarity 96.5%; Pred. No. 0;
	Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;
Oy	1 CAAGCTTCTTAATTTTCGTG--CTCCTCTTACACGCCCAATCGGTTTTGCACCT 58
Dd	
	1 CAAGCTCTTAAATTTCCTTGCTCTCTCTCTTACACGGCCAATGGTGTTCGCACCT 60
Oy	59 TTGAGGCCCTCAATCCAGACATTTCTATAAGCATATTGCAAGAGGGGCTCTTAATTG 118
Dd	
	61 TTCAGGCCCTCAATCAACAACATCTCATATAAGCATATTGCAAGAGGGGCTCTTAATTG 120
Oy	119 TTGCATTGAGTTTATCGCTATAGCTAGAGGAATTTCTAATTTAGGGAGGCTCTAGAATT 178
Dd	
	121 TTGCATGAGTTGAACCAATATAGACTAGAGGAATTTCAATTAGGGAGGCTCTAGAATT 180
Oy	179 TGCACTAACCTTCATAATGCGCTCTGTGAGCTGTGAGCTGATTAATGAACAAGATGTCTAG 238
Dd	
	181 TGCACTAACCTTCATAATGAGCTCTGAGAGCTGTGATGATTTGATTAACAAGATGTCTAG 240
Oy	239 GCAGAAATGTGCGCGAAATCTCTCACCGGAGGAAGTATTTCAACTGATGAGATCTCTT 298
Dd	
	241 GCAGAAATGTGCGCGAAATCTCTCACCGGAGGAAGTATTTCAACTGATGAGATCTCTT 300
Oy	299 GATPATATGTAAACCTGTTGCAGTATATAAATCATCTTACCTGCGCTCTAGGCAACC 358
Dd	
	301 GATPATATGTAAACCTGTTGCAGTATATAAATCATCTTACCTGCGCTCTAGGCAACC 360
Oy	359 ATCGTTTCTTCCAAGATGCTTGAACCTACMAAATTGGAGCAAAAGGCCAAAAGAAAGTCAAG 418
Dd	
	361 ATCGTTTCTTCCAAGATGCTTGAACCTACMAAATTGGAGCAAAAGGCCAAAAGAAAGTCAAG 420
Oy	419 ATCTACTGGGATGAGTACTGTTTCAACTATAAAGGATGTTAATAACATTTACAGAAAACTGA 478
Dd	
	421 ATCTACTGGGATGAGTACTGTTTCAACTATAAAGGATGTTAATAACATTTACAGAAAACTGA 480
Oy	479 AGTTAGGAGAGATTGTTCTTCCATATTTGCTCATGCTATGCTAGTGTAGCTTCAAGGGCT 538
Dd	
	481 AGTTAGGAGAGATTGTTCTTCCATATTTGCTCATGCTATGCTAGTGTAGCTTCAAGGGCT 540
Oy	539 GCAATTTCAATTGAATTCATCTCATGATTTATTTGAATTGAGTTCAAGCTTTTCGAGA 598
Dd	
	541 GCAATTTCAATTGAATTCATCTCATGATTTATTTGAATTGAGTTCAAGCTTTTCGAGA 600
Oy	599 ATTACAGACATTPATGTTTCTGAAAACTAATTCCTCATATTTTGGAGAAAGAGAG 658
Dd	
	601 ATTACAGACATTPATGTTTCTGAAAACTAATTCCTCATATTTTGGAGAAAGAGAG 660
Oy	659 TGATGAGATTAATTTGAGCCCTCTCTCTGCTGCAAMCCTGTAGCGGAGACAAG 718
Dd	
	661 TGATGAGATTAATTTGAGCCCTCTCTCTGCTGCAAMCCTGTAGCGGAGACAAG 720

QY	719	AGGTGGCGAATATACACACAGAGACTAAAGTAGCTTTTACCGTTGGATCCACG	778
Db	721	AGGTGGCGAATATACACACAGAGACTAAAGTAGCTTTTACCGTTGGATCCACG	780
QY	779	TTTAACTAATGGCAGAGAAATGGAATACCCCTACTTATGATGGAACCGGGTTAG	838
Db	781	TTTAACTAATGGCAGAGAAATGGAATGATCCCTGCTGATGATGGAACCGGGTTAG	840
QY	839	ATATCCGAGGCAACAGAGCTTGGCGCAATTGAGATGATACCAATTCACACAGC	898
Db	841	ATATCCGAGGCAACAGAGCTTGGCGCAATTGAGATGATACCAATTCACACAGC	900
QY	899	CATAGCCACACTCTCTCTGAGACCGTGGTCTAAAGTTATTTAGACAGCAGCTGTGCT	958
Db	901	CATAGCCACACTCTCTCTGAGACCGTGGTCTAAAGTTATTTAGACAGCAGCTGTGCT	960
QY	959	CCGTGCTACTAAGACAGAAAGTTATCTGCTGAGGATCAGAGGCTAGAACGCACCTCT	1018
Db	961	CCGTGCTACTAAGACAGAAAGTTATCTGCTGAGGATCAGAGGCTAGAACGCACCTCT	1020
QY	1019	TCTTCAGAAACGCCAATCTATCATCTTCACAGAGTCCAGCCAAATGCGCTTGAGCAGT	1078
Db	1021	TCTTCAGAAACGCCAATCTATCATCTTCACAGAGTCCAGCCAAATGCGCTTGAGCAGT	1080
QY	1079	AATCTCTACCGGGGTATGACGAGATGGAAGTCGATGACATGTGGAGATTTGAAAGATG	1138
Db	1081	AATCTCTACCGGGGTATGACGAGATGGAAGTCGATGACATGTGGAGATTTGAAAGATG	1140
QY	1139	CCAGATGCTTATGACTCTTGTGATGTGTAATTAAGATGTAAGAAAGCAATTCATGACATCTT	1199
Db	1141	CCAGATGCTTATGACTCTTGTGATGTGTAATTAAGATGTAAGAAAGCAATTCATGACATCTT	1200
QY	1199	GAACTCGTTTATGAAACAAAGGCTATAGCAGATGCGATATCTCTGGCGATGGA	1258
Db	1201	GAACTCGTTTATGAAACAAAGGCTATATGCAATGTGATATCTCTGGCGATGGA	1260
QY	1259	AGCATTTTCAAGATTTTACGAGAAAGGTGACACCGTTATCTATCTCTCTGGTGTG	1318
Db	1261	AGCATTTTCAAGATTTTACGAGAAAGGTGACATGATCTATCTCTCTGGTGTG	1320
QY	1319	GAGATTTGTTTGTATTAACATATGAGAACCATGACCTTGGACCTCAGCCACATGACAA	1378
Db	1321	GAGATTTGTTTGTATTAACATATGAGAACCATGACCTTGGACCTCAGCCACATGACAA	1380
QY	1379	CTGCAATACCATCTCGAGAAATGCCGTAATAGTCTAGACACCAACCCACACACACAA	1438
Db	1381	CTGCAATACCATCTCGAGAAATGCCGTAATAGTCTAGACACCAACCCACACACACAA	1434
QY	1439	CAACAGTGTGATGCTCCACAGTGCCTCAACACCAACCAACATTAACATTTGGATCAACC	1498
Db	1435	CAACAGTGTGATGCTCCACAGTGCCTCAACACCAACCAACATTAACATTTGGATCAACC	1494
QY	1499	CAATGACATTAACAACAAGACATGTGTGACACACAGACATTAACAGACAGACAAAGT	1558
Db	1495	CAATGACATTAACAACAAGACATGTGTGACACACAGACATTAACAGACAGACAAAGT	1555
QY	1559	AATTAATATGAGAAATCTCCGGCTTTATATGATACCGATTTTCGATGTGTAATTAATCT	1618
Db	1554	AATTAATATGAGAAATCTCCGGCTTTATATGATACCGATTTTCGATGTGTAATTAATCT	1613
QY	1619	TCTTTCTCT-AAAAAATGTTTAAAGACCAACAATTTTATATGTTATGTTATTCACAT	1677
Db	1614	TCTTTCTCTAAAAAATGTTTAAAGACCAACAAGATTTTATTTGTTATGTTATTCACAT	1673
QY	1678	GATTACATTTTATGTTAAAAAANAAT 1705	
Db	1674	GATTACATTTTATGTTAAAAAANAAT 1701	

AA447755;
 16-NOV-2000 (first entry)
 Aberrantly spliced VRN2 nucleole acid.
 Vernalization gene: VRN2; plant characteristic: flowering time;
 leaf size; leaf shape; shade avoidance response; reproduction;
 breeding; pollination; cultivation; ss.
 Arabidopsis thaliana var 'Columbia'.
 W0200044918-A1.
 03-AUG-2000.
 28-JAN-2000; 2000WC-GB00248.
 28-JAN-1999; 99GB-0001927.
 (PLAN-) PLANT BIOSCIENCE LTD.
 Dean C, Gendall A;
 WPI: 2000-499333/44.
 P-PSDB; AAB00062.
 Isolated vernalization gene VRN2 is used to produce transgenic plants
 with altered vernalization response, flowering time, leaf size and/or
 shape or shade avoidance response for maximized reproductive success
 Claim 8; Page 74-75; 105bp; English.
 Isolated nucleole acid sequences obtained from the VRN2 locus of a
 plant encode polypeptides which are capable of affecting one or more
 vernalization responses such as, flowering time, leaf size and/or
 shape or the shade avoidance response of a plant into which the
 nucleole acid is introduced. Introducing such sequences into plants
 to alter these characteristics maximises the reproductive success of
 the plant. This cDNA was generated from an aberrantly spliced VRN2
 gene.
 Sequence 1/37 BP; 543 A; 330 C; 359 G; 505 T; 0 other;

RESULT 5
AAA47755
ID AAA47755 ctctcgcacgca cduw: 1777 bp

Query Match	89.3%	Score 1537.6	DB 21	Length 1737
Best Local Similarity	95.3%	Pred. No. 0		
Matches 1649	Conservative	0	Mismatches 49	Indels 32
			Gaps	5
Qy	1	CAAGCTTCTCAATTTTGCTG--CTCTCTTACACAGCCATCGGTGTTTCGAGCT	58	
Db	1	CAAGCTTCTCAATTTTGCTGCTCTCTCTCTTACAGGCCATCGGTGTTTCGAGCT	60	
Qy	59	TTGAGGCCCTCATCCACAGACATTTGTATATAGCATATTGAGAAAGGGGGTTCTAATG	118	
Db	61	TTGAGGCCCTCATACAGACATTTGTATATAGCATATTGAGAAAGGGGGTTCTAATG	120	
Qy	119	TTTGCAATGAGTTATCCGTATAGAGCTAGAGAAATTCATATTAGGGGAGCCTCAGAGT	178	
Db	121	TTTGCAATGAGTTGAAACATATAGACGTAGAGAAATTCATATTAGGGGAGCCTCAGAGT	180	
Qy	179	TGACACAACTTATATATGGCTCTTGAGCGTTGTGAGTGAATTAAGAAAGAAATGTGAG	238	
Db	181	TGACACAACTTATATATCAGCTCTGGAGCTGTGATATGATTTTGACACAGAAATGTGAG	240	
Qy	239	GCAGAAATTGTGCGCGGAAATCCTACCGGAGAGAGTATTTCAACTGATGAGATCTCTT	298	
Db	241	GCAGAAATTGTGCGCGGAAATCCTACCGGAGAGAGTATTTCAACTGATGAGATCTCTT	300	
Qy	299	GATATATTTGTAACCTGTTGACATATATACATCTTCACTCTGCTCTTAGGCAACC	358	
Db	301	GATATATTTGTAACCTGTTGACATATATACATCTTCACTCTGCTCTTAGGCAACC	360	

OY	359	ATCGTTTCTCCAGAGTCCTGAGCTACGAAATTTGGAGCAAGCGCCAAAGAAAGTCAAG	418
OY	361	ATCGTTTCTCCAGAGTCCTGAGCTACGAAATTTGGAGCAAGCGCCAAAGAAAGTCAAG	420
OY	419	ATCTACTGGAGTGTAGTTTCAACTAATAAGATTGTAAATAACATATACAGAAACCTGA	478
Db	421	ATCTACTGGAGTGTAGTTTCAACTAATAAGATTGTAAATAACATATACAGAAACCTGA	480
OY	479	AGTTAGGAGGAGATTGTTCTTGTCCATATTTGCTATAGCTATAGGTAGCTTCA-----	531
Db	481	AGTTAGGAGGAGATTGTTCTTGTCCATATTTGCTATAGCTATAGGTAGCTTCAAGGTGG	540
OY	532	-----AGGGGCTGCAATTTCAATTGGAATTCATCTCAGATTTATTTGAAT	576
Db	541	CAACTATTTACAGTACGAGGGGCTGGAATTTCAATTGGAATTCATCTCAGATTTATTTGAAT	600
OY	577	TTGAGTTCAACGCTTTTGCAGAAATATCCAGACAGTTAATGTTCTGTAAACCTTAATTCCT	636
Db	601	TTGAGTTCAACGCTTTTGCAGAAATATCCAGACAGTTAATGTTCTGTAAACCTTAATTCCT	660
OY	637	TCATATTTTGAGGAAGAAGAGTATGACGATAAATTTGAGCCCTTCTCTCTGCTCGA	696
Db	661	TCATATTTTGAGGAAGAAGAGTATGATGATTAATTTGAGCCCTTCTCTCTGCTCGA	720
OY	697	AACCTCGTAACGGGAGACAAAGAGGTGGCAGAAATTAACACAGAGACTTAAATAGTCT	756
Db	721	AACCTCGTAACGGGAGACAAAGAGGTGGCAGAAATTAACACAGAGACTTAAATAGTCT	780
OY	757	TTTACCGGTTGGATTACACAGTTTAACATATAGGACAGAAATAGGAATGCAACCTACTTA	816
Db	781	TTTACCGGTTGGATTACACAGTTTAACATATAGGACAGAAATAGGAATGCAACCTACTTA	840
OY	817	ATGATGGAACCGTGTTTAGAGATATCCGAGGCAACAGAGCTTCTGAGCAATTTGAGA	876
Db	841	ATGATGGAACCGTGTTTAGAGATATCCGAGGCAACAGAGCTTCTGAGCAATTTGAGA	900
OY	877	TGACCAGACATTTCCACACAGCCATAGCCCACTCTTCTGTGAGCGTGTGCTAAAGTTA	936
Db	901	TGACTTAGAACATTTCCACACAGCCATAGCCCACTCTTCTGTGAGCGTGTGCTAAAGTTA	960
OY	937	TATTGACAAAGGGAAGCTGTGTCCTCGTACTAAGACAAAGAAAGTTATCTGAGGCAT	996
Db	961	TATTGACAAAGGGAAGCTGTGTCCTCGTACTAAGACAAAGAAAGTTATCTGAGGCAT	1020
OY	997	CAGAGGCTAGAAAGCCACTACTTCTTTCGAAAGCCCAATTCATTCATCTCAGAGAGCC	1056
Db	1021	CAGAGGCTAGAAAGCCACTACTTCTTTCGAAAGCCCAATTCATTCATCTCAGAGAGCC	1080
OY	1057	AGCCAAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGAGTAAAGTCATGACG	1116
Db	1081	AGCCAAATGGCGCTTGAGCAAGTAATGTCTGACCGGGATAGCGAGAGTAAAGTCATGACG	1140
OY	1117	ATGTGTGAGATTTTGAAGATGCGCAAGATGCTGATAGACTTGTGAGATGGAATTAAGATG	1176
Db	1141	ATGTGTGAGATTTTGAAGATGCGCAAGATGCTGATAGACTTGTGAGATGGAATTAAGATG	1200
OY	1177	AAAAGCAATTCATCATCTTTTGAAGACGCTTTGTAAGAAAACAAGGGTTTATACAGATG	1236
Db	1201	AAAAGCAATTCATCATCTTTTGAAGACGCTTTGTAAGAAAACAAGGGTTTATACAGATG	1260
OY	1237	GTCATATTCCTTTGGGCAATGTGACAGATTTTACGAGAAAGATGTGACCGTT	1296
Db	1261	GTCATATTCCTTTGGGCAATGTGACAGATTTTACGAGAAAGATGTGACCGTT	1320
OY	1297	ACATATACCTCTTGTGAGTGTGGAGATTGTTTGAATTAACATATGCAACCATGACCTTG	1356
Db	1321	ACATATACCTCTTGTGAGTGTGGAGATTGTTTGAATTAACATATGCAACCATGACCTTG	1380
OY	1357	TCGACTAGAGCCACATCAACAATGCAATACATCTCGAGAAATTTGCCGTAAATAGCTCAG	1416
Db	1381	TCGACTAGAGCCACATCAACAATGCAATACATCTCGAGAAATTTGCCGTAAATAGCTCAG	1440
OY	1417	ACACCCACACCAACACACACAGTGTGATGCTCCACGTGACTCAAAACACCAACA	1476

[illegible]

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142827.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145092.
PR 23-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145324.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149375.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154379.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

Fri Jun 20 09:03:38 2003

us-09-890-220-1.1mg

PR 28-OCT-1999: 9905-0161992.
 PR 28-OCT-1999: 9905-0161993.
 PR 29-OCT-1999: 9905-0162142.

56.9%; Score 979.8; DB 21; Length 1497;

Best Local Similarity 89.7%; Pred. No. 1.3e-232; Indels 102; Gaps 2;

Matches 1148; Conservative 0; Mismatches 27;

Query Match
 355 ACCCATGCTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGCCCAAAAGAAAGT 414
 320 AGCCATCGTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGCCCAAAAGAAAGT 379
 415 CAGATCTACTGGAGTGTAGTGTTCCTTCACTAAGATTTGATTAACATTTACAGAAAA 474
 380 CAGATCTACTGGAGTGTAGTGTTCCTTCACTAAGATTTGATTAACATTTACAGAAAA 439
 475 CTGAAGTTAGGAGAGATGTTCTTCTGCAATTTGCTATGCTATGCTAGTCTTCAAG 534
 440 CTGAAGTTAGGAGAGATGTTCTTCTGCAATTTGCTATGCTATGCTAGTCTTCAAG 499
 535 GGCATCATTTTCAATTTGATTTCAATTCATTCATGATTTATTTGATTTGATTTGATTTG 594
 500 GGCATCATTTTCAATTTGATTTCAATTCATTCATGATTTATTTGATTTGATTTGATTTG 559
 595 AAGATACAGACAGATTTATGTTCTGTAATACTTATTCATTTGATTTGATTTGATTTG 654
 560 CTGTAAATCCAACTAA----- 577
 655 GAAGTATGATGATTAATTTGAGCCCTTCTCTCTGCTGCAAACTCGTAAGCCGAGAC 714
 578 -----TGTTCACACTGCAAACTCGTAAGCCGAGAC 610
 715 AAGAGGTGACAAATTAACACAGAGACCTTAAGTATGCTTTTACCGTGGATTCAC 774
 611 AAGAGGTGACAAATTAACACAGAGACCTTAAGTATGCTTTTACCGTGGATTCAC 670
 775 CCAGTTTAACTAATGAGCAGAAATGGAATCACCTTAACTTAAGTGAAGAACCTGGTT 730
 671 CCAGTTTAACTAATGAGCAGAAATGGAATCACCTTAACTTAAGTGAAGAACCTGGTT 730
 835 TAGATATCCGAGAGCAAGAGCTGCTGCAATTTGATTTGATTTGATTTGATTTGATTTG 894
 731 TAGATATCCGAGAGCAAGAGCTGCTGCAATTTGATTTGATTTGATTTGATTTGATTTG 790
 895 CAGCCATAGCCACTCTCTCTGAGCGTGTCTTAAGTATTTGATTTGATTTGATTTGATTTG 954
 791 CAGCCATAGCCACTCTCTCTGAGCGTGTCTTAAGTATTTGATTTGATTTGATTTGATTTG 850
 955 TGGTCCCTGCTAATGAGCAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 910
 851 TGGTCCCTGCTAATGAGCAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1041
 1008 -----AGCCACTCTCTCTGAGCGTGTCTTAAGTATTTGATTTGATTTGATTTGATTTG 970
 911 CCATATTAACGAGAAATTTCTTCCAGCCACTCTCTCTTCAAGAAAGCCCAATTTATTC 970
 1042 ATTCTCAGAGAGTCCAGCAATGCGCTTGAGCAATTTGATTTGATTTGATTTGATTTGATTTG 1101
 971 ATTCTCAGAGAGTCCAGCAATGCGCTTGAGCAATTTGATTTGATTTGATTTGATTTGATTTG 1030
 1102 ATGAAGTATGAGAGTGTGAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1161
 1031 ATGAAGTATGAGAGTGTGAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1090
 1162 ATGTGAATTAAGATGAAGAAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1221
 1091 ATGTGAATTAAGATGAAGAAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1150
 1222 GGGTTATGAGAGTGTGAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1281
 1151 GGGTTATGAGAGTGTGAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1210
 1282 AAGAGTTGACACCGTTACTGATCTCTGCTGCTGTTGAGATTTGATTTGATTTGATTTGATTTG 1341

DB 1211 AAGATTTGACCGTGTACTGATCACTCTTGTGGTGTGAGATTTGATTTGATTTGATTTGATTTG 1270
 1342 GGAACCATGAGACTGTGTGACTACAGCCATCAATCAACATGCAATTCATTCATTCATTCATTCAT 1401
 1271 GGAACCATGAGACTGTGTGACTACAGCCATCAATCAACATGCAATTCATTCATTCATTCATTCAT 1330
 1402 GCGGTATATGCTCAGAC 1461
 1331 GCGGTATATGCTCAGAC 1390
 1462 ACTCAAC 1521
 1391 ACTCAAC 1450
 1522 ATGTTGAC 1568
 1451 ATGTTGAC 1497

RESULT 7

AAA47752 standard; DNA; 6338 BP.

AAA47752;

16-NOV-2000 (first entry)

VRN2 nucleic acid.

Vernalization gene; VRN2; plant characteristic; flowering time;

leaf size; leaf shape; shade avoidance response; reproduction;

breeding; pollination; cultivation; ds.

Arabidopsis thaliana var 'landsberg erecta'.

WO200044918-A1.

03-AUG-2000.

28-JAN-2000; 2000WO-GB00248.

28-JAN-1999; 99GB-0001927.

(PLANT-) PLANT BIOSCIENCE LTD.

Dean C, Gendall A;

WPI: 2000-499333/44.

Isolated vernalization gene VRN2 is used to produce transgenic plants

with altered vernalization response, flowering time, leaf size and/or

shape or shade avoidance response for maximized reproductive success

Claim 9; Page 68-70; 105pp; English.

Isolated nucleic acid sequences obtained from the VRN2 locus of a

plant encode polypeptides which are capable of affecting one or more

vernalization responses such as flowering time, leaf size and/or

shape or shade avoidance response of a plant into which the

nucleic acid is introduced, introducing such sequences into plants

to alter these characteristics maximises the reproductive success of

the plant.

Sequence 6338 BP; 1932 A; 1076 C; 1134 G; 2193 T; 3 other;

Query Match 23.9%; Score 412.2; DB 21; Length 6338; Best Local Similarity 98.8%; Pred. No. 9.4e-92; Indels 0; Gaps 0;

Matches 414; Conservative 1; Mismatches 4;

1304 ACTCTGCTGCTGAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1363
 5102 AATTTCAAGGTGTGAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 5161

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149802.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 18.2%; Score 314.2; DB 21; Length 319;
 Best Local Similarity 99.1%; Pred. No. 5.4e-68;
 Matches 316; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1375 ACAACGCAATACCATCTCCGAGATTCGCCGTAATAGCTCAGACACGACGACGACACA 1434
 DB 1 ACAACGCAATACCATCTCCGAGATTCGCCGTAATAGCTCAGACACGACGACGACACA 60
 OY 1435 ACAACGAGTGTGATCGTCGAGTACTCAACGACGACGACGACGACGACGACGACGAC 1494
 DB 61 ACAACGAGTGTGATCGTCGAGTACTCAACGACGACGACGACGACGACGACGACGAC 120
 OY 1495 ATCCGATGATTAAC 1554
 DB 121 ATCCGATGATTAAC 180
 OY 1555 AAGTAATTAATAGGAAATCTCGGCTTTATGATACGATTTATCGATTTGACTTA 1614
 DB 181 AAGTAATTAATAGGAAATCTCGGCTTTATGATACGATTTATCGATTTGACTTA 240
 OY 1615 TTCTCTCTCTTAAATAATGTTAGACAAACAATTTTATATGTATGTTAGTATCA 1674
 DB 241 TTCTCTCTCTTAAATAATGTTAGACAAACAATTTTATATGTATGTTAGTATCA 300
 OY 1675 ACTGATTACATTTTACTT 1693
 DB 301 ACTGATTACATTTTACTT 319

RESULT 10

ANAC4147
 ID ANAC4147 standard; DNA: 640 BP.
 AC ANAC4147;
 AC ANAC4147;
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41806.
 XX Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX
 XX 06-SEP-2000.
 PD
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130591.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134321.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134841.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 01-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137522.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140961.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.

```
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144337.
PR 20-JUL-1999; 99US-0144338.
PR 21-JUL-1999; 99US-0144339.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145090.
PR 22-JUL-1999; 99US-0145091.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145228.
PR 23-JUL-1999; 99US-0145229.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145277.
PR 27-JUL-1999; 99US-0145278.
PR 27-JUL-1999; 99US-0145279.
PR 28-JUL-1999; 99US-0145280.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146387.
PR 03-AUG-1999; 99US-0146388.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
```

```
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 17.8%; Score 306.4; DB 21; Length 640;
Best Local Similarity 98.8%; Pred. No. 6,1e-66;
Matches 318; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1373 CAACACTGCAATACATCTCGAGATTCGCGTAATAGCTCAGACACACACACCAA 1432
    |||||||
DB 319 CAACACTGCAATACATCTCGAGATTCGCGTAATAGCTCAGACACACACACCAA 378
    |||||||
QY 1433 CAACAACAAGTGTGTCGTCGAGTCAACACCAACAAATATATGTTGGA 1492
    |||||||
DB 379 CAACAACAAGTGTGTCGTCGAGTCAACACCAACAAATATATGTTGGA 438
    |||||||
QY 1493 TCATCCCAATGACATTAACAACAAGACATGTTGACACACAGACATTAACAGAG- 1551
    |||||||
DB 439 TCATCCCAATGACATTAACAACAAGACATGTTGACACACAGACATTAACAGAG- 498
    |||||||
QY 1552 ACAAGTATTAATAGAAATCTCGGCTTTATGATACGATTTATCGATTGTAAC 1611
    |||||||
DB 499 ACAAGTATTAATAGAAATCTCGGCTTTATGATACGATTTATCGATTGTAAC 558
    |||||||
QY 1612 TTATTCCTTCTTAAATAATGTTAGAGCAACAAATTTTATGTTAGTGTAT 1671
    |||||||
DB 559 TTATTCCTTCTTAAATAATGTTAGAGCAACAAATTTTATGTTAGTGTAT 618
    |||||||
QY 1672 TCAACTGATTAATTTAGTT 1693
    |||||||
DB 619 TCAACTGATTAATTTAGTT 640
    |||||||

RESULT 11
AAZ36947
ID AAZ36947 standard; cDNA: 2280 BP.
XX
AC AAZ36947;
DT 13-MAR-2000 (first entry)
DE cDNA encoding a MPCL protein having flowering regulating activity.
KW Flowering regulating activity; MPCL; flowering; germination;
KW super early flowering mutation; altered flowering time;
KW flowering regulating gene; food crop; vegetable; flowering inhibition;
KW productivity; ss.
XX
```


PR 24-JUN-1999; 99JP-0179043.
 XX (MITA) MITSUI CHEM INC.
 XX Yoshiida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;
 DR WPI; 2000-064612/06.
 DR P-PSDB; AAV53933.
 XX
 PT Novel DNA used to produce transgenic plants with altered floral
 PT regulation which can have increased crop yields -
 XX
 PS Claim 6; Page 40-44; 53pp; English.
 CC The present sequence encodes a protein having a flowering regulating
 CC activity, which is designated Os-MPC1. An Arabidopsis MPC1 is also
 CC disclosed in the specification. The rice and Arabidopsis cDNAs show
 CC significant homology with each other. A naturally occurring mutation
 CC of the MPC1 gene eliminates normal flowering regulating ability of
 CC plants, and leads to flowering immediately after germination (super
 CC early flowering mutation). The MPC1 polynucleotide sequence can be
 CC used to produce plants with altered flowering times in comparison
 CC with wild type plants, by enhancing or inhibiting the expression of
 CC the flowering regulating gene. Antisense polynucleotides can be used
 CC to reproduce the effects of the mutated MPC1 gene. This alteration
 CC can be used to increase the yield of food crops. Flowering inhibition
 CC of vegetables increases their productivity.
 XX
 SO Sequence 2248 BP; 641 A; 454 C; 493 G; 660 T; 0 other;

Query Match 14.7%; Score 252.4; DB 21; Length 2248;
 Best Local Similarity 56.2%; Pred. No. 2.4e-52;
 Matches 555; Conservative 0; Mismatches 376; Indels 57; Gaps 2;

OY 424 CTGGAGTGGTGTTCACATTAAGATTGATTAACACATTACAGAAACTGAGTTA 483
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 939 CTGGCAATGTCTTTTAACTACACAACTACAAATACATGCAAAAAACCGAGTCA 998
 OY 484 GGGAGGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 999 CTGAAGATTTTCTGGCCATTTTCTGTCGACCAAGTGGCGCTTGAAGGCTTAGAT 1058
 OY 544 TTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAATTTGAATTTGAATTTGA 603
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1059 GTCACTTAAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1118
 OY 604 AGACAGTTAATGTTCTGTAATACTTAATTCCTCATATTTGAG---GAAGAAGAGAG 660
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1119 AGCGCTTAATGTTAGTCTGAAGACTGATTTCTTGAGAACAGAGCTTTTGCTGAG 1178
 OY 661 ATGAGATTAATTTGAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1179 TTGATCCAGACATTAACATTTTCTGACCGCTCAAGATTTAAGAGGGTAAAGGGTGG 1238
 OY 721 GTGGCAGAAATTAACACGAGACTTAAAGTATGCTTTTACCGTTGATTCACCCAGT 780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1239 AAATCTCAAGATTAATTAAGCATGTACATCCACATATGTGATTCAGGATCACTG 1298
 OY 781 TAACTAATGACAGCAAAATGATCACCCTAATTAATGATGAGAAACCGTGTAGAT 840
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1299 AAGATGCCAGCAGAGATCTCAAGACGATTAAGTCAAGAGGAAATAGTGTGTTGAG 1358
 OY 841 ATCCGAGAGCAAGAGCTGTGCTGCAAAATTTGACATGACCAAGCAATTCACAGCA 900
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1359 CACACGCTTCTGTTGATCTGCTAATTAATTAACAGGTAACATCTTCAGACAC- 1413
 OY 901 TAGCCCACTTCTCTGACGCTGCTGCTAAGATTATTTGACAAAGCAAGCAAGCTGCT 960
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1414 -----AACAGTGTAC 1424
 OY 961 CTGCTACTAAGCAAGAAATTAATCTGCTGACGATCAAGAGCTGAAGCAACCTACTTC 1020
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1425 AGTTGGAGACAGAAAGAGCTGCTGTTGAACGAGCTATCCAGAAATGCGCAGCTCC 1484

OY 1021 TTCAGAAAGCCCAATTCATATCATCTCACAGAGTCAGCCAAATGGCGTTGAGCAAGTAA 1080
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1485 TACAAAAGCCAGTCTTTCTTCTTCACAGGCGCTCAACCAATGGCATGAGCAAGATT 1544
 OY 1081 TGTCTGACCGGATGAGCAGAGATGAAGTCGATGACGATGTTGACAGATTTTGAAGATCGCC 1140
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1545 TCTCAGATCGTATGATGTAAGATGAAGTTGATGATGATGATGATGATGATGATGATG 1604
 OY 1141 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1605 GAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1664
 OY 1201 ACTGCTGTAAGAAACAAAGGTTATAGCATGATGATGATGATGATGATGATGATGATG 1260
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1665 ATTCATTTGTTGGAAACAAAGGCTACATGCGATGCGATATTTCTCGGCGATCGAG 1724
 OY 1261 CATTTTCAAGATTTTACGAGAAAGATGACCGCTTACTCATCTCTTCTGCTGTTGA 1320
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1725 CATCTCGCAGTTTCATGAGCAAGAACTGTACAAATCCACCTCTACTATAGCTTGA 1784
 OY 1321 GATGTTTGTGTTAACTATGAGACCATGACCTGTGCTGACCTACCATCAACACT 1380
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1785 GGTGTTTATGCTCAAACTCTGGAACACAGCTCTACTGATGCGGAGCCATGAATGCT 1844
 OY 1381 GCAATTCATCTCGAAGATTCGCTAA 1408
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1845 GCAACCAATTTCTTGAAGGCTACTGTA 1872

RESULT 13

AAA47758
 ID AAA47758 standard; cDNA; 558 BP.

AAA47758;
 16-NOV-2000 (first entry)

At Hyp 2245035 (ATFC7_4) (modified) cDNA.

Vernalization gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ss.

Arabidopsis thaliana.

WO20044918-A1.

03-AUG-2000.

28-JAN-2000; 2000WO-GB00248.

28-JAN-1999; 99GB-0001927.

(PLAN-) PLANT BIOSCIENCE LTD.

Dean C, Gendall A;

WPI; 2000-499333/44.

P-PSDB; AAB00065.

Isolated vernalization gene VRN2 is used to produce transgenic plants

with altered vernalization response, flowering time, leaf size and/or

shape or shade avoidance response for maximized reproductive success

Disclosure; Page 76; 105pp; English.

Isolated nucleic acid sequences obtained from the VRN2 locus of a
 plant encode polypeptides which are capable of affecting one or more
 vernalization responses such as, flowering time, leaf size and/or
 shape or the shade avoidance response of a plant into which the
 nucleic acid is introduced. Introducing such sequences into plants
 to alter these characteristics maximises the reproductive success of


```

PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.
PR 23-JUL-1999: 99US-0145218.
PR 23-JUL-1999: 99US-0145224.
PR 26-JUL-1999: 99US-0145276.
PR 27-JUL-1999: 99US-0145813.
PR 27-JUL-1999: 99US-0145818.
PR 27-JUL-1999: 99US-0145919.
PR 28-JUL-1999: 99US-0145951.
PR 02-AUG-1999: 99US-0146386.
PR 02-AUG-1999: 99US-0146388.
PR 03-AUG-1999: 99US-0147038.
PR 04-AUG-1999: 99US-0147204.
PR 04-AUG-1999: 99US-0147302.
PR 05-AUG-1999: 99US-0147192.
PR 05-AUG-1999: 99US-0147260.
PR 06-AUG-1999: 99US-0147303.
PR 06-AUG-1999: 99US-0147416.
PR 09-AUG-1999: 99US-0147493.
PR 10-AUG-1999: 99US-0147935.
PR 11-AUG-1999: 99US-0148171.
PR 12-AUG-1999: 99US-0148319.
PR 13-AUG-1999: 99US-0148341.
PR 13-AUG-1999: 99US-0148565.
PR 16-AUG-1999: 99US-0148684.
PR 17-AUG-1999: 99US-0149175.
PR 18-AUG-1999: 99US-0149426.
PR 20-AUG-1999: 99US-0149722.
PR 20-AUG-1999: 99US-0149723.
PR 20-AUG-1999: 99US-0149929.
PR 23-AUG-1999: 99US-0149902.
PR 23-AUG-1999: 99US-0149930.
PR 25-AUG-1999: 99US-0150566.
PR 26-AUG-1999: 99US-0150884.
PR 27-AUG-1999: 99US-0151065.
PR 27-AUG-1999: 99US-0151066.
PR 27-AUG-1999: 99US-0151080.
PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.

```

```

PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161892.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

```

Query Match

9.9%; Score 170.8; DB 21; Length 525;

Best Local Similarity 96.2%; Pred. No. 2.2e-32; Mismatches 7; Indels 0; Gaps 0;

Matches 175; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 355 ACCCATCGTTTCTCCAGATGCTTGAAGTACAAATGAGCAAGCCCAAGAAAGT 414
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 326 AGCCATCGTTTCTGCCAAGATGCTTGAAGTACAAATGAGCAAGCCCAAGAAAGT 385

QY 415 CAAGATCTACTGGGATGATGTTTCAACTATTAAGATTGTATTAACATTTACAAAA 474
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 386 CAAGATCTACTGGGATGATGTTTCAACTATTAAGATTGTATTAACATTTACAAAA 445

QY 475 CTGAAGTTAGGAGGATGTTCTGTCCATTTTGTCTAAGCATGATGAGTCAAGG 534
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 446 CTGAAGTTAGGAGGATGTTCTGTCCATTTTGTCTAAGCATGATGAGTCAAGG 505

QY 535 GG 536
    |
Db 506 TG 507

RESULT 15
AAAA7757
ID AAAA7757 standard; cDNA; 325 BP.
XX
AC AAAA7757;
XX
DT 16-NOV-2000 (first entry)
XX
DE A1163743 Expressed sequence tag (modified) cDNA.
XX
KW Vernalization gene; VRN2; plant characteristic; flowering time; poplar;
KW leaf size; leaf shape; shade avoidance response; reproduction; EST;
KW breeding; pollination; cultivation; expressed sequence tag; ss.
XX
OS Populus tremula.
XX
PN MO200044918-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-GB00248.
XX

```

```
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI: 2000-499333/44.
DR F-PSDB: AAB00064.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
XX PS Disclosure: Page 76; 105pp; English.
XX
XX CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA represents expressed sequence tag clone A1153743
CC from poplar trees (Populus tremula x Populus tremuloides) which has
CC homology in the C-terminal region of the Arabidopsis thaliana VRN2 gene.
XX
SQ Sequence 325 BP; 102 A; 60 C; 76 G; 87 T; 0 other;

Query Match 9.5%; Score 163.8; DB 21; Length 325;
Best Local Similarity 76.4%; Pred. No. 9.8e-31;
Matches 201; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 951 GCTGTGTCCTGCTACTAAGACAGAAAGTTATCTGTCAGCATCAGAGCTAGAGC 1010
   || || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 GCCATGCTACAAATTTGCAAGACAGAAATTTATCAATGTAGCGTCTGACATGAGAAC 121

QY 1011 CACCTACTTCTTCAGAAACGCCAATTCATCTCAGAGTCAGCCCAATGGCGCT 1070
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 CGTACACTCTCCACAAACGACATTTTTCACCTACATAGAGCTCAGCCATGGCAGCT 181

QY 1071 GAGCAATTAATGTCTGACCGGATAGCGAGATGAAGTCGATGATGTTCAGATTTT 1130
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 GAGCAAGTTATGTGATGATGCGGATAGTGAAGTGAAGTGAAGTGTTCAGATTTT 241

QY 1131 GAGATGCCAGATGCTGATGACTTTGTGGATGAATAAGATGAAAAGCAATTCATG 1190
   ||||| || | | | | | | | | | | | | | | | | | | | | | | | |
Db 242 GAAAGCCGAAAGATGCTTGATTTGTAGACCTGAAGTGAAGTGAAGCAAAATGATG 301

QY 1191 CATCTTTGGAACCTGTTTGTAG 1213
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 302 CACTTGTGGAACCTCATTTGTGAG 324
```

Search completed: June 19, 2003, 20:10:38
Job time : 430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 19:56:48 ; Search time 90 Seconds

(without alignments)
5867.749 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctctcaatttgct.....aatgattctgtataact 1722

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCtus.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	69	4.0	198	PCT-US95-10668-3	Sequence 3, Appli
2	69	4.0	198	PCT-US95-10668-4	Sequence 4, Appli
3	65.8	3.8	198	PCT-US95-10668-1	Sequence 1, Appli
4	65.8	3.8	198	PCT-US95-10668-2	Sequence 2, Appli
5	62	3.6	5511	US-08-928-361B-2	Sequence 1, Appli
6	62	3.6	7334	US-08-928-361B-1	Sequence 1, Appli
7	60.4	3.5	5163	US-08-700-651-1	Sequence 1, Appli
8	60.4	3.5	5163	US-08-928-361B-4	Sequence 4, Appli
9	60.4	3.5	5318	US-08-700-651-2	Sequence 1, Appli
10	60.4	3.5	5318	US-08-928-361B-3	Sequence 3, Appli
11	59.2	3.4	1086	US-08-415-751-27	Sequence 27, Appli
12	59.2	3.4	1086	US-08-415-751-28	Sequence 28, Appli
13	53.4	3.1	1690	US-08-276-452A-24	Sequence 24, Appli
14	53.4	3.1	1690	US-08-798-744-24	Sequence 24, Appli
15	51	3.0	2032	US-09-241-581B-5	Sequence 5, Appli
16	51	3.0	2032	US-08-265-428-5	Sequence 5, Appli
17	51	3.0	2032	PCT-US95-07721-5	Sequence 5, Appli
18	49.8	2.9	1086	US-08-415-751-47	Sequence 47, Appli
19	48.2	2.8	688	US-08-998-416-915	Sequence 915, App
20	47.6	2.8	1430	US-08-276-452A-25	Sequence 25, Appli
21	47.6	2.8	1430	US-08-798-744-25	Sequence 25, Appli
22	45.8	2.7	2190	US-09-625-188-19	Sequence 19, Appli
23	43.6	2.5	1107	US-07-814-964-12	Sequence 12, Appli
24	43.4	2.5	3292	US-08-258-442-12	Sequence 12, Appli
25	43.4	2.5	3292	US-08-328-809-7	Sequence 7, Appli
26	43.4	2.5	3292	PCT-US92-11107-12	Sequence 12, Appli
27	43.4	2.5	3292		

C 28	40.8	2.4	87350	3	US-08-781-891-79	Sequence 79, Appli
C 29	40.8	2.4	87543	4	US-09-791-211-3	Sequence 3, Appli
C 30	39.8	2.3	915	4	US-09-134-001C-1741	Sequence 1741, Ap
C 31	39.6	2.3	1529	2	US-08-477-451-17	Sequence 17, Appli
C 32	39.6	2.3	1529	2	US-08-477-451-21	Sequence 21, Appli
C 33	39.6	2.3	2295	1	US-08-375-300-3	Sequence 3, Appli
C 34	39.6	2.3	2295	3	US-09-177-431-3	Sequence 3, Appli
C 35	39.6	2.3	2295	5	PCT-US95-16930-3	Sequence 3, Appli
C 36	39.6	2.3	4080	1	US-08-375-300-1	Sequence 1, Appli
C 37	39.6	2.3	4080	3	US-09-177-431-1	Sequence 1, Appli
C 38	39.6	2.3	4080	5	PCT-US95-16930-1	Sequence 1, Appli
C 39	38.8	2.3	658	4	US-08-998-416-595	Sequence 595, App
C 40	38.8	2.3	912	4	US-09-134-001C-1352	Sequence 1352, Ap
C 41	38.6	2.2	399	3	US-08-713-559-6	Sequence 6, Appli
C 42	38.4	2.2	444	4	US-09-178-509-2	Sequence 2, Appli
C 43	38.4	2.2	450	1	US-08-090-523-28	Sequence 28, Appli
C 44	38.4	2.2	450	1	US-08-398-627-28	Sequence 28, Appli
C 45	38.4	2.2	450	1	US-08-406-857-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
PCT-US95-10668-3
Sequence 3, Application PC/TUS9510668
GENERAL INFORMATION:
APPLICANT: James Eberwine
TITLE OF INVENTION: A Method of Sequencing Proteins by
TITLE OF INVENTION: Epitope Ordering and Protein
TITLE OF INVENTION: Restriction Mapping
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10668
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/294,133
FILING DATE: August 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 198
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: No
PCT-US95-10668-3
Query Match 4.0%; Score 69; DB 5; Length 198;
Best Local Similarity 60.3%; Pred. NO. 1.2e-09;
Matches 114; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
1373 CAACAACGCAATGCATCCGAGAAATGGCGTAATAGTCACACACACCACCA 1432
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rn1

Page 2

[illegible]

```

1 RESULT 2
2 PCT-US95-10668-4
3 : Sequence 4, Application PC/TUS9510668
4 : GENERAL INFORMATION:
5 :   APPLICANT: James Eberwine
6 :   TITLE OF INVENTION: A Method of Sequencing Proteins by
7 :   TITLE OF INVENTION: Epitope Ordering and Protein
8 :   TITLE OF INVENTION: Restriction Mapping
9 :   NUMBER OF SEQUENCES: 4
10 :   CORRESPONDENCE ADDRESS:
11 :     ADDRESSEE: Jane Massey Licata, Esq.
12 :     STREET: 210 Lake Drive East, Suite 201
13 :     CITY: Cherry Hill
14 :     STATE: NJ
15 :     COUNTRY: USA
16 :     ZIP: 08002
17 :   COMPUTER READABLE FORM:
18 :     MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
19 :     COMPUTER: IBM 486
20 :     OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
21 :     SOFTWARE: WORDPERFECT 5.1
22 :   CURRENT APPLICATION DATA:
23 :     APPLICATION NUMBER: PCT/US95/10668
24 :     FILING DATE: Herewith
25 :   CLASSIFICATION:
26 :     PRIOR APPLICATION DATA:
27 :       APPLICATION NUMBER: 08/294,133
28 :       FILING DATE: August 22, 1994
29 :     ATTORNEY/AGENT INFORMATION:
30 :       NAME: Jane Massey Licata
31 :       REGISTRATION NUMBER: 32,257
32 :       REFERENCE/DOCKET NUMBER: PENN-0137
33 :       TELECOMMUNICATION INFORMATION:
34 :         TELEPHONE: (609) 779-2400
35 :         TELEFAX: (609) 779-8488
36 :         INFORMATION FOR SEQ. ID NO.: 4:
37 :           SEQUENCE CHARACTERISTICS:
38 :             LENGTH: 196
39 :             TYPE: Nucleic Acid
40 :             STRANDEDNESS: Single
41 :             TOPOLOGY: Linear
42 :           ANTI-SENSE: NO
43 : PCT-US95-10668-4

```

[illegible]

Db	121	CACACACACACACACACACACACACACACACACACACAA	180
Q7	1553	CAAACTAAT	1561
Db	181	CAAGCTGAT	189

RESULT 3
 PCT-US95-10668-1
 Sequence 1, Application PC/TUS9510668
 GENERAL INFORMATION:
 APPLICANT: James Eberwine
 TITLE OF INVENTION: A Method of Sequencing Proteins by
 TITLE OF INVENTION: Epitope Ordering and Protein
 TITLE OF INVENTION: Restriction Mapping
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jane Massey Licata, Esq.
 STREET: 210 Lake Drive East, Suite 201
 CITY: Cherry Hill
 STATE: NJ
 COUNTRY: USA
 ZIP: 08002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 COMPUTER: IBM 486
 OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10668
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/294,133
 FILING DATE: August 22, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: PENN-0137
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-2400
 TELEFAX: (609) 779-8488
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 198
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 ANTI-SENSE: No
 PCT-US95-10668-1

[illegible]

```

PCT-US95-10668-2
; Sequence 2, Application PC/TUS9510668
; GENERAL INFORMATION:
; APPLICANT: James Eberwine
; TITLE OF INVENTION: A Method of Sequencing Proteins by
; TITLE OF INVENTION: Epitope Ordering and Protein
; TITLE OF INVENTION: Restriction Mapping
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10668
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/294,133
; FILING DATE: August 22, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: PENN-0137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
; PCT-US95-10668-2

Query Match          3.8%; Score 65.8; DB 5; Length 198;
Best Local Similarity 59.3%; Pred. No. 8.9e-09;
Matches 112; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 1373 CACAACATGCAATACATCTCGAGAAATTCGGTAATAGTCAGACACCACCAACCA 1432
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CACAACAACAACAACAACAACAACAACAACAAGATCCCAACAACAACAACAACA 60

QY 1433 CACAACAAGAGTGTGATGTCGCCAGTGACTCAAAACCAACAACAATTAAGATTGTGA 1492
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 120

QY 1493 TCATCCCATGATATAACAACAACAAGAAATTTGACACCAAGCAATTAAGAGAGAGA 1552
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 180

QY 1553 CAAAGTAAT 1561
    ||| ||| |||
Db 181 CAAAGTGAT 189

RESULT 5
US-08-928-361B-2
; Sequence 2, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

```

```

; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, HANA
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5511 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-928-361B-2

Query Match          3.6%; Score 62; DB 3; Length 5511;
Best Local Similarity 57.1%; Pred. No. 5.2e-07;
Matches 113; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1358 CGACTCAGCCACCATCAACATGCAATACATCCGCAATTCGCAATTCGCGTAATAGCTCAGA 1417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2150 CTACCTAAGGCAACAACAACAACAACAATCTTAATCCATCTTACAAACAACACTCAA 2209

QY 1418 CACCACACCACCAACAACAACAACAAGATGATCGTCCAGTGAACCAACCA 1477
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2210 AACCAACAACAACAACAACAACAAGGTTCCAGTAAGCCACCAATAGCCACAACAACA 2269

QY 1478 CAATAACATTTGTGATCATCCCATGACATAAACAACAAGAACATGTTGACAACAAGA 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2270 CAACTATTAAGGCAATAGTTACAAACAACAACAACAAGCAACAACAACAACA 2329

QY 1538 CAATAACAGCAGAGACAA 1555
    ||| ||| |||
Db 2330 CAGTGCACAGCAACTA 2347

RESULT 6
US-08-928-361B-1
; Sequence 1, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto

```

STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk.
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcclin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verry, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480,76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7334 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
OS-08-928-361B-1

```

Query Match: 3.6%; Score 62; DB 3; Length 7334;
Best Local Similarity 57.1%; Pred. No. 6e-07;
Matches 113; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

OY      1358  CCACATCAGCCACCATCAACAATGCGAATACCATCTCCGAGATTGCCGTAAATAGTCTAGA   1417
          ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB       3818  CTACTTAAGGCACCAACAACAACCAAAATTCCTTTATTCATATCTTTACCAACAACATCTAAA   3877
OY      1418  CACCACCAACCAACAACAACAAGAGTGATGGTCCAGTGCACTGAACCAACAACA   1477
          |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB       3878  AACCAACAACAACAACAACAAGAGTGTCAGGTAAAGCCAATAAGCCAACAACA   3937
OY      1478  CAATACATTTGTGATCATCCCAATACATTTAAACAACAAGAATGTGGCAACAAGGA   1537
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB       3938  CAACATTTAACCAATAGTTTACAAACAACAACAACAAGAAAGCAACAACAACAACA   3997
OY      1538  CATTAACAGCACAGACAACA   1555
          || | || | || | || |
DB       3998  CAGTGCACAACACACACTRA   4015

RESULT 7
US-08-700-651-1
: Sequence 1, Application US/08700651B
: Patent No. 6015882
: GENERAL INFORMATION:
: APPLICANT: PETERSEN, CAROLYN
: APPLICANT: LEECH, JAMES
: APPLICANT: NELSON, RICHARD, C.
: APPLICANT: CHU, JIKI
: TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
: TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
: FILE REFERENCE: 480.19-4(HV)
: CURRENT APPLICATION NUMBER: US/08/700,651B
: EARLIER FILING DATE: 1997-08-14
: EARLIER APPLICATION NUMBER: 08/415,751
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 5163

```

Query Match	Best Local Similarity	Score	DB 3:	Length
Matches 112; Conservative	56.68;	3.58;	0;	5163;
			Mismatches 86;	Gaps 0;
Qy	1358	CGACTCAGCCACCATATACACTGCATACACTCCTCGAGATTCGCCGTATATGCTCACA	1417	
Db	1799	TACTATAGGCGAACACACACACACACAACTCTTATATCCATCATTTACACACACAGCTCACA	1858	
Qy	1418	CACACACACACACACACACACACAAACAGTGTGGTGTGCTGCCAGTACTCAACACACACAA	1477	
Db	1859	AACCCACACACACACACACACACAAAGAGTCTCCAGTAGCGACCAATGTGCCACACACACAA	1918	
Qy	1478	CATATACCATTTGTGGATCATCCCAATGACATTAACACACAGAACATTTGTGACACACAGA	1557	
Db	1919	CACACTTAAAGCCCAATATGTTTACCAACACACACACAAACAAAGCAACACACACACACAA	1978	
Qy	1538	CATATACAGCAGAGACAA	1555	
Db	1979	CAGTCCCAACGACCACTA	1996	

```

US-08-928-361B-4
: Sequence 4, Application us/08928361B
: Patent No. 6071518
:
: GENERAL INFORMATION:
: APPLICANT: Petersen, Carolyn
: TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
: TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS AND FRAGMENTS
: TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
: TITLE OF INVENTION: SPECIES INFECTIONS
: NUMBER OF SEQUENCES: 30
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PETERS, VERNY, JONES & BIRKA
: STREET: 385 Sherman Avenue, Suite 6
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306-1840
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/928,361B
: FILING DATE: 12-SEP-1997
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/026,062
: FILING DATE: 13-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Verny, Hana
: REGISTRATION NUMBER: 30,518
: REFERENCE/DOCKET NUMBER: 480,76-1(HV)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-1678
:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5163 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
:
: MOLECULE TYPE: DNA (genomic)
:
: US-08-928-361B-4

```

Matches 112: Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
OY 1358 CGACTCAGCCACATCAACATGCAATACCATCTCGAGATGGCCGTAAATAGCTGGA 1417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 CTACTAAGGCAACAACAACCACTCTTAATCCATATTAACAACAACACTCAA 1858
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1418 CACCACGACCAACAACAACAAGTGGATGCTCCAGTGCATCAACACCAACA 1477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 AACCAACAACAACAACAAGGTTCCAGGTAAGCCACCAATAGCCACACACAA 1918
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1478 CAATTAACATTTGGATGATCCCAATGACATTAACAACAACAATGTTGACACAGGA 1537
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CAACATTAAGCCATAGTTACTACAACAACAACAAGCAACAACAACAACAACA 1978
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1538 CAATTAACGACGAGACAA 1555
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 CAGTGCCACGACAACTA 1996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 9

```
US-08-700-651-2
: Sequence 2, Application US/08700651B
: Patent No. 6015882
: GENERAL INFORMATION:
: APPLICANT: PETERSEN, CAROLYN
: APPLICANT: LEECH, JAMES
: APPLICANT: NELSON, RICHARD, C.
: APPLICANT: GUT, JIRI
: TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
: TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF cryptosporidium parvum
: FILE REFERENCE: 480.19-4(HV)
: CURRENT APPLICATION NUMBER: US/08/700,651B
: EARLIER FILING DATE: 1997-08-14
: EARLIER APPLICATION NUMBER: 08/415,751
: EARLIER FILING DATE: 1995-04-03
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 2
: LENGTH: 5318
: TYPE: DNA
: ORGANISM: Cryptosporidium parvum
US-08-700-651-2
```

Query Match 3.5%: Score 60.4; DB 3: Length 5318;

Best Local Similarity 56.6%: Pred. No. 1.4e-06;

Matches 112: Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
OY 1358 CGACTCAGCCACATCAACATGCAATACCATCTCGAGATGGCCGTAAATAGCTGGA 1417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 CTACTAAGGCAACAACAACCACTCTTAATCCATATTAACAACAACACTCAA 1858
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1418 CACCACGACCAACAACAACAAGTGGATGCTCCAGTGCATCAACACCAACA 1477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 AACCAACAACAACAACAAGGTTCCAGGTAAGCCACCAATAGCCACACACAA 1918
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1478 CAATTAACATTTGGATGATCCCAATGACATTAACAACAACAATGTTGACACAGGA 1537
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CAACATTAAGCCATAGTTACTACAACAACAACAAGCAACAACAACAACAACA 1978
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1538 CAATTAACGACGAGACAA 1555
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 CAGTGCCACGACAACTA 1996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 10

```
US-08-928-361B-3
: Sequence 3, Application US/08928361B
: Patent No. 6071518
: GENERAL INFORMATION:
: APPLICANT: PETERSEN, CAROLYN
: APPLICANT: NELSON, RICHARD, C.
: APPLICANT: GUT, JIRI
: TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
: TITLE OF INVENTION: PPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
: TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
```

TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

TITLE OF INVENTION: SPECIES INFECTIONS

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESS: PETERS, VERNY, JONES & BIRSA

STREET: 385 Sherman Avenue, Suite 6

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-1840

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,361B

FILING DATE: 12-SEP-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,062

FILING DATE: 13-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: VERNY, Hana

REGISTRATION NUMBER: 30,518

REFERENCE/DOCKET NUMBER: 480.76-1(HV)

TELEPHONE: 650-324-1677

TELEFAX: 650-324-1678

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5318 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-928-361B-3

Query Match 3.5%: Score 60.4; DB 3: Length 5318;

Best Local Similarity 56.6%: Pred. No. 1.4e-06;

Matches 112: Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
OY 1358 CGACTCAGCCACATCAACATGCAATACCATCTCGAGATGGCCGTAAATAGCTGGA 1417
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 CTACTAAGGCAACAACAACCACTCTTAATCCATATTAACAACAACACTCAA 1858
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1418 CACCACGACCAACAACAACAAGTGGATGCTCCAGTGCATCAACACCAACA 1477
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 AACCAACAACAACAACAAGGTTCCAGGTAAGCCACCAATAGCCACACACAA 1918
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1478 CAATTAACATTTGGATGATCCCAATGACATTAACAACAACAATGTTGACACAGGA 1537
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CAACATTAAGCCATAGTTACTACAACAACAACAAGCAACAACAACAACAACA 1978
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1538 CAATTAACGACGAGACAA 1555
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 CAGTGCCACGACAACTA 1996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

```
US-08-415-751-27
: Sequence 27, Application US/08415751
: Patent No. 5643772
: GENERAL INFORMATION:
: APPLICANT: PETERSEN, CAROLYN
: APPLICANT: LEECH, JAMES
: APPLICANT: NELSON, RICHARD, C.
: APPLICANT: GUT, JIRI
: TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
: TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
: TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
: TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
: TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
```

;; TITLE OF INVENTION: DIAGNOSIS AND KIT
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: PHILLIPS, MOORE, LEMPLO & FINLEY
;; STREET: 385 Sherman Avenue, Suite 6
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94306-1840
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
;; COMPUTER: PC
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/415,751
;; FILING DATE: 03-Apr-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/071,880
;; FILING DATE: June 1, 1993
;; APPLICATION NUMBER: 07/891,301
;; FILING DATE: May 29, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hana Dolezalova
;; REGISTRATION NUMBER: 30,518
;; REFERENCE/DOCKET NUMBER: 480.19-2 (HND)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-1677
;; TELEFAX: (415) 324-1678
;; INFORMATION FOR SEQ ID NO: 27:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1086 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Cryptosporidium parvum
US-08-415-751-27

Query Match
Best Local Similarity 56.8%; Score 59.2; DB 1; Length 1086;
Matches 109; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 1376 CAACTGCAATACCATCTCTGGAATGGCCGTAATAGTCTCAGACACCACCACCACCAACAA 1435
DB 341 CAACTGCAATACCATCTCTGGAATGGCCGTAATAGTCTCAGACACCACCACCACCAACAA 400
QY 1436 CAACACAGTGTGGATGCTCCAGTGAATCAACCAACCAACCAATATGATGATCA 1495
DB 401 CAACACAGTGTGGATGCTCCAGTGAATCAACCAACCAACCAATATGATGATCA 460
QY 1496 TCCCAATGACATTAACCAACCAACCAATGATGACCAACCAACCAATTAACAGAGACAA 1555
DB 461 AGCCAATAGTGAACCAACCAACCAATGATGACCAACCAACCAATTAACAGAGACAA 520
QY 1556 AGTAATTAATA 1567
DB 521 CGACAACACTACTA 532

RESULT 12
US-08-415-751-28/c
; Sequence 28, Application US/08415751
; Patent No. 5643772
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
; TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA

;; TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
;; TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
;; TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS AND KIT
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: PHILLIPS, MOORE, LEMPLO & FINLEY
;; STREET: 385 Sherman Avenue, Suite 6
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94306-1840
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
;; COMPUTER: PC
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/415,751
;; FILING DATE: 03-Apr-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/071,880
;; FILING DATE: June 1, 1993
;; APPLICATION NUMBER: 07/891,301
;; FILING DATE: May 29, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hana Dolezalova
;; REGISTRATION NUMBER: 30,518
;; REFERENCE/DOCKET NUMBER: 480.19-2 (HND)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-1677
;; TELEFAX: (415) 324-1678
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1086 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Cryptosporidium parvum
US-08-415-751-28

Query Match
Best Local Similarity 56.8%; Score 59.2; DB 1; Length 1086;
Matches 109; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 1376 CAACTGCAATACCATCTCTGGAATGGCCGTAATAGTCTCAGACACCACCACCACCAACAA 1435
DB 746 CAACTGCAATACCATCTCTGGAATGGCCGTAATAGTCTCAGACACCACCACCACCAACAA 687
QY 1436 CAACACAGTGTGGATGCTCCAGTGAATCAACCAACCAACCAATATGATGATCA 1495
DB 686 CAACACAGTGTGGATGCTCCAGTGAATCAACCAACCAACCAATATGATGATCA 627
QY 1496 TCCCAATGACATTAACCAACCAACCAATGATGACCAACCAACCAATTAACAGAGACAA 1555
DB 626 AGCCAATAGTGAACCAACCAACCAATGATGACCAACCAACCAATTAACAGAGACAA 567
QY 1556 AGTAATTAATA 1567
DB 566 CGACAACACTACTA 555

RESULT 13
US-08-276-452A-24
; Sequence 24, Application US/08276452A
; Patent No. 5646029
; GENERAL INFORMATION:
; APPLICANT: Chen, Chao-Guang
; APPLICANT: Mao, Shao-Lim
; APPLICANT: DU, He

Query Match	3.1%	Score 53.4;	DB 1;	Length 1690;
Best Local Similarity	55.7%;	Pred. NO. 6.4e-05;		
Matches 102; Conservative	0;	Mismatches 81;	Indels 0;	Gaps 0;

```

RESULT 14
US-08-798-744-24
: Sequence 24: Application US/08798744
: Patent No. 5830747
: GENERAL INFORMATION:
: APPLICANT: Chen, Chao-Guang
: APPLICANT: Mau, Shao-Lim
: APPLICANT: Du, He
: APPLICANT: Gane, Allison M
: APPLICANT: Bacic, Antony
: APPLICANT: Clarke, Adrienne E
: TITLE OF INVENTION: Plant Arabidnagalactan Protein (AGP) Genes
: NUMBER OF SEQUENCES: 91
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Greenlee and Winner, P.C.
: STREET: 5370 Manhattan Circle, Suite 201
: CITY: Boulder
: STATE: Colorado
: COUNTRY: United States of America
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/798,744
: FILING DATE: 13-FEB-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/276,452
: FILING DATE: 18-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Caruthers, Jennie M.
: REGISTRATION NUMBER: 34,464
: REFERENCE//DOCKET NUMBER: 27-91A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303)499-8080
: TELEFAX: (303)499-8089
: TELEX: 49617824
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1690 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 60..1442
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 1..38
: OTHER INFORMATION: /note= "Nucleotide sequence
: OTHER INFORMATION: obtained by PCR which does not overlap with the
: OTHER INFORMATION: cDNA clone"

```

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rni

Page 8

FEATURE:
NAME/KEY: misc.feature
LOCATION: 60-128
OTHER INFORMATION: /note- "Predicted transmembrane
segment"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 135-179
OTHER INFORMATION: /note- "Derived amino acid sequence
corresponding to the peptide sequence by protein
microsequencing"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 135-179
OTHER INFORMATION: /note- "Amino acids 27 to 36, 38,
and 40 are identical to that in the peptide
obtained by direct microsequencing"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 135-179
OTHER INFORMATION: /note- "Amino acid 26 may also be
undetermined residues"
US-08-798-744-24

Query Match 3.1%; Score 53.4; DB 2; Length 1690;
Best Local Similarity 55.7%; Pred. No. 6.4e-05;
Matches 102; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 1373 CAACAACCTGCAATACATCCGAGATTCGCTATAGCTCAGACACACACACCA 1432
DB 920 CCAGATTCATGAAACAACAATGCTTCGAGAGTTACAACAACAACAACA 979
QY 1433 CAACAACAAGTGTGATGCTCCAGTCACTCAACAACAACAACAACAACA 1492
DB 980 CAACAACAAGTGTGATGCTTCGAGATTCACAACAACAACAACAACAACA 1039
QY 1493 TCATCCATGACATAAACAACAACAATGTCACAACAACAACAACAACA 1552
DB 1040 CTCGAGATTCACAACAACAACAATGCTTCACAGAGATTCACAACAACA 1099
QY 1553 CAA 1555
DB 1100 CAA 1102

RESULT 15
US-09-241-581B-5
Sequence 5, Application US/09241581B
Patent No. 6350859

GENERAL INFORMATION:

APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: Class BI and CI Scavenger Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,581B
FILING DATE: 02-Feb-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: MIT6620
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2032 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc.feature
LOCATION: 40-1926
OTHER INFORMATION: /Function = "Nucleotides 40 through
1926 encode the amino acid sequence
for the Drosophila Melanogaster
Scavenger Receptor Class CI."
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-241-581B-5

Query Match 3.0%; Score 51; DB 4; Length 2032;
Best Local Similarity 57.4%; Pred. No. 0.00032;
Matches 112; Conservative 0; Mismatches 80; Indels 3; Gaps 1;
QY 1358 GCACTGACCCACATCAACAACATGCAATACCATCTCGAGATTCGCTATAGCTCAGA 1417
DB 1346 CAACAAGGCGCAACTCAAAAGCAACAACAACAACAACAACAACAACA 1405
QY 1418 CACGACGACGACCAACAACAACAAGTGTGATGCTCCAGTCACTCAACAACAACA 1477
DB 1406 CAACGCGAAGCCACACACACGAC---TCAACACACCAAGCTCAACTCTACAA 1462
QY 1478 CAATACATTTGTGATCATCCATGACATTAACAACAACAACAATGTTGACAACAAGA 1537
DB 1463 CGCTACACTTCAACAACAACAACAACAACAACAACAACAACAACAACA 1522
QY 1538 CAATACAGCAGAGA 1552
DB 1523 AAACAACAATAATGA 1537

Search completed: June 19, 2003, 22:10:18
Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 21:25:52 ; Search time 285 Seconds
(without alignments)
8866.349 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctctcaattgcttgc.....aatgattcgttataact 1722

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCRT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187.2	10.9	2556	9 US-09-938-842A-2418	Sequence 2418, App
2	92.4	5.4	888	9 US-09-938-842A-453	Sequence 453, App
3	55.8	3.2	2886	10 US-09-801-368-131	Sequence 131, App
4	54.8	3.2	4985	9 US-10-094-240-10	Sequence 10, Appl
5	52	3.0	1236	9 US-10-077-584-3	Sequence 3, Appl
6	51.6	3.0	867	10 US-09-216-393-340	Sequence 340, App
7	51.6	3.0	867	10 US-09-216-393-342	Sequence 342, App
8	51.6	3.0	1397	10 US-09-216-393-345	Sequence 345, App
9	51.6	3.0	1397	10 US-09-216-393-345	Sequence 345, App
10	51.4	3.0	333	9 US-09-938-842A-1152	Sequence 3152, App
11	45.8	2.7	2901	10 US-09-801-368-371	Sequence 371, App
12	45.2	2.6	423	9 US-09-918-995-7147	Sequence 7147, App
13	44.8	2.6	2277	10 US-10-026-188-6	Sequence 223, App
14	44	2.6	155074	9 US-10-026-188-6	Sequence 6, Appl
15	43.4	2.5	1560	9 US-09-893-519A-82	Sequence 82, Appl
16	43	2.5	1257	10 US-09-822-830A-428	Sequence 428, App
17	43	2.5	2217	9 US-09-874-162A-6	Sequence 6, Appl
18	43	2.5	2328	9 US-09-874-162A-9	Sequence 9, Appl
19	43	2.5	2481	10 US-09-764-864-124	Sequence 124, App

20	43	2.5	3150	9 US-09-938-842A-243	Sequence 243, App
21	43	2.5	3545	10 US-09-764-864-144	Sequence 144, App
22	43	2.5	4409	9 US-09-874-162A-7	Sequence 7, Appl
23	43	2.5	4441	9 US-09-874-162A-4	Sequence 4, Appl
24	43	2.5	4441	10 US-09-969-347-169	Sequence 169, App
25	43	2.5	6030	9 US-10-239-676-164	Sequence 164, App
26	42.6	2.5	456	10 US-09-864-761-4249	Sequence 4249, App
27	42.2	2.5	58985	9 US-09-901-152-3	Sequence 3, Appl
28	41.8	2.4	500	9 US-09-991-936-531	Sequence 521, App
29	41.8	2.4	5085	9 US-10-198-846-9854	Sequence 9854, App
30	41.4	2.4	493	10 US-09-864-761-5863	Sequence 5863, App
31	41.2	2.4	14649	9 US-10-239-676-122	Sequence 122, App
32	41	2.4	1530	10 US-09-759-777-113	Sequence 113, App
33	40.8	2.4	779	10 US-09-772-134B-69	Sequence 69, Appl
34	40.8	2.4	2614	9 US-09-822-846-491	Sequence 491, Appl
35	40.8	2.4	6167	9 US-10-239-676-44	Sequence 42, Appl
36	40.8	2.4	6604	10 US-09-880-107-1748	Sequence 1748, App
37	40.8	2.4	335913	9 US-09-754-853A-2	Sequence 2, Appl
38	40.6	2.4	335913	9 US-09-754-853A-3	Sequence 3, Appl
39	40.6	2.4	1075	10 US-09-864-761-19241	Sequence 19241, A
40	40.6	2.4	1403	10 US-09-864-761-2513	Sequence 2513, App
41	40.4	2.3	2031	10 US-09-801-368-301	Sequence 301, App
42	40.4	2.3	1830121	9 US-10-329-960-1	Sequence 1, Appl
43	40.2	2.3	411	10 US-09-960-352-14521	Sequence 14521, A
44	40.2	2.3	17419	9 US-10-239-676-100	Sequence 100, App
45	40	2.3	6306	9 US-10-239-676-130	Sequence 130, App

ALIGNMENTS

RESULT 1

US-09-938-842A-2418
Sequence 2418, Application US/09938842A

Patent No. US20020160378A1
GENERAL INFORMATION:

APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPT000-3
CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 2418
LENGTH: 2556

TYPE: DNA
ORGANISM: Arabidopsis thaliana

US-09-938-842A-2418

Query Match 10.9%; Score 187.2; DB 9; Length 2556;
Best Local Similarity 89.7%; Pred. No. 2.8e-37;
Matches 201; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

816 AATGATGAACCGGTTTGGATATCCCGGACAGAGCTTGGGCAATTGAG 875
|||||
2332 ACTGAGGAAACCGGTGTTTGGATATCCCGGACAGAGCTTGGGCAATTGAG 2391
|||||
876 ATGACCAACCAATTCACACAGCCATTCCTTGGAGCGTGTCTAAATT 935
|||||
2392 ATGACTACCAACATTCACACAGCCATTCCTTGGAGCGTGTCTAAATT 2451
|||||
936 ATATTGACCAACGAGCTGTGTCTCTCTACTAGACAGAAATTATCTGAGGCA 995
|||||
2452 ATATTGACCAACGAGCTGTGTCTCTCTACTAGACAGAAATTATCTGAGGCA 2511
|||||

QY 996 TCAGAGGCTAGAACCCACCCTACTTCTTCGAAACGCCAATTCTA 1039
 |||||
 Db 2512 TCAGAGGCTAGAAAGTTTGTTCATCATGACACCCCGTCATCATA 2555

RESULT 2

US-09-938-842A-453
Sequence 453, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krepes, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 453
LENGTH: 888
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-453

Query Match	5.4%;	Score 92.4;	DB 9;	Length 888;
Best Local Similarity	73.4%;	Pred. No. 2.4e-13;		
Matches 160;	Conservative	0;	Mismatches 1;	Indels 57;
				Gaps 1;

OY	1062	ATGGGCGCTAGCAATATATGCTCTACCCGGATAGGAGAGATGAAGTCGATGACGATGTT	1122
Db	1	ATGGCGCTTAGCAAGTATATGCTCTATCGGATATGGAGAGATGAAGTCGATGACGATGTT	60
OY	1122	GCAGATTTTAAATATCCGACG-----	1144
Db	61	GCAGATTTTAAATATCCGAGATATCCATGATTTCTTCTGCTTCAATTAACTAGACAA	120
OY	1143	-----ATGCTTGATGACTTTTGAGATGATGAATATGAATGAAGAAACAA	1184
Db	121	CAGAAAATGTAATGACATATGCTTGATGACTTTTGAGATGATGAATATGAATGAAGAAACAA	180
OY	1185	TTTCATGATCTTTGGAACTCGTTTGTAAAGAAAACAAG	1222
Db	181	TTTCATGATCTTTGAAATCGTTTGTAAAGAAAACAAG	218

RESULT 3

```

US-09-801-368-131
: Sequence 131, Application US/09801368
: Patent No. US20020128250A1
: GENERAL INFORMATION:
: APPLICANT: Busby, Robert
: APPLICANT: Cali, Brian
: APPLICANT: Hecht, Peter
: APPLICANT: Holtzman, Doug
: APPLICANT: Madden, Kevin
: APPLICANT: Maxon, Mary
: APPLICANT: Milne, Todd
: APPLICANT: NO. US20020128250A1man, Thea
: APPLICANT: Royer, John
: APPLICANT: Salama, Sofie
: APPLICANT: Sherman, Amir
: APPLICANT: Silva, Jeff
: APPLICANT: Summers, Eric
: TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

```

```

? FILE REFERENCE: 109272.147
? CURRENT APPLICATION NUMBER: US/09/801,368
? CURRENT FILING DATE: 2001-03-07
? PRIOR APPLICATION NUMBER: US 09/487,558
? PRIOR FILING DATE: 2000-01-19
? PRIOR APPLICATION NUMBER: US 60/160,587
? PRIOR FILING DATE: 1999-10-20
? NUMBER OF SEQ ID NOS: 440
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 131
? LENGTH: 2886
? TYPE: DNA
? ORGANISM: Saccharomyces cerevisiae
? US-09-801-368-131

```

Query Match	3.2%;	Score 55.8;	DB 10;	Length 2886;
Best Local Similarity	54.0%;	Pred. No. 0.0011;		
Matches 114;	Conservative 0;	Mismatches 97;	Indels 0;	Gaps 0;

Oy	1366	ACCACCATTAACAACTGGCAATPACCATCTGTGAAATPCCCTAATTAAGTACGACACAC	1423
Db	1461	AGCGATCATCAACACAAACAGTAACAAAGACGACGACAAACACAAATTAACGATTA	1520
Oy	1424	CACCAACAAACAAACAACAGTGTGATGCTGCCAGTGCCTAAACACCAACCAATTA	1489
Db	1521	CGATTACGATTAACAAATTAATAGCAATTAATTAATTAATTAATTAATTAATTAATTA	1580
Oy	1484	CATTGTGATTAATCCCAATGATATTAACACACAGAAACATGTGTGACAAAGGACAAATA	1544
Db	1581	TATATACATTAATTAATTAACAAATTAACAAACAAACAAACAAATTAATTAATTAATTA	1640
Oy	1544	CAGCAGACCAAAAGTATTAATTAAGAAAAAT	1574
Db	1641	TATTAATAATACGTTGATTAACAATAAATAGT	1671

RESULT 4

US-10-094-240-10/c
Sequence 10, Application US/10094240
Publication No. US20030082637A1
GENERAL INFORMATION:
APPLICANT: ZWITZEL, LAURENCE J.
TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
FILE REFERENCE: N8389
CURRENT APPLICATION NUMBER: US/10/094, 240
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 10/056,405
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: 60/264,649
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 4985
TYPE: DNA
ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match	3.28;	Score 54.8;	DB 9;	Length 4985;
Best Local Similarity	47.68;	Pred. No. 0.0028;		
Matches 161; Conservative	0;	Mismatches 177;	Indels 0;	Gaps 0;

[illegible]

```

QY 1563 AATAGAAATCTCCGCTTTATGATACCGATTATCGATGTAATTATCTCT 1622
    ||| ||| | | | | | | | | | | | | | | | | | | | | | |
DB 2883 TAATTAATTAAGATAGTAATTAATTAATTAATTAATTAATTAATTA 2824
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1623 TCTTAATAAATGTTAGAGCAACAATTTTATAGTGTATTCAGTGA 1682
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 2823 TAATTAATTAATTAATTAATTAAGATAATTAATTAATTAATTAAT 2764
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1683 CATTTTAGTTAAAAAATGATCTGCTTAA 1720
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 2763 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2726
    ||| ||| | | | | | | | | | | | | | | | | | | | | |

```

RESULT 5

```

US-10-077-584-3
; Sequence 3, Application US/10077584
; Publication No. US20030073610A1
; GENERAL INFORMATION:
; APPLICANT: LINDQUIST, SUSAN
; APPLICANT: KROBITSCH, SYLVIA
; APPLICANT: OUTIERO, TIAGO F.
; TITLE OF INVENTION: YEAST SCREENS FOR THE TREATMENT OF HUMAN DISEASE
; FILE REFERENCE: ARCD:36705
; CURRENT APPLICATION NUMBER: US/10/077,584
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269,157
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(513)
US-10-077-584-3

```

Query Match

```

Best Local Similarity 3.0%; Score 52; DB 9; Length 1236;
Matches 103; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY 1367 CACATCAACAACCTGATATCCATGAGATATGCGTATATAGTCAGACACGAC 1426
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 58 CAGCAACAGCAACACAGCAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 117
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1427 CACCAACAACAACAGTGTGATCGTCCAGTCTCAACACCAACCAACATACAT 1486
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 118 CAGCAACAACAGCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 177
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1487 TGTGATATCCCATGATGATTAACAACAACAACATTTGACAAACAGGACATACAG 1546
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 178 CAACAGCAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 237
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1547 CAGAGACA 1554
    ||| |||
DB 238 CAGCAACA 245

```

RESULT 6

```

US-09-216-393-340
; Sequence 340, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDI PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19

```

```

; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 340
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(867)
US-09-216-393-340

```

Query Match

```

Best Local Similarity 3.0%; Score 51.6; DB 10; Length 867;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```

```

QY 1370 CATCAACAACCTGATATCCATGAGATATGCGTATATAGTCAGACACGACAC 1429
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 485 CATCAACTACCACTACTACACTACACACTACACTACACTACACTACACTAC 544
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1430 CAACAACAACAACAGTGTGATCGTCCAGTGTCAACAACAACAACAATTAATGT 1489
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 545 CACCAACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 604
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1490 GGATCATCCCATGATGATTAACAACAACAACAATTTGACAAACAGGACATACAGAG 1549
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 605 CACCAACAACAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 664
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1550 AGACAA 1555
    ||| |||
DB 665 CGCCAA 670

```

RESULT 7

```

US-09-216-393-342
; Sequence 342, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDI PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 342
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
US-09-216-393-342

```

Query Match

```

Best Local Similarity 3.0%; Score 51.6; DB 10; Length 867;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

```

```

QY 1370 CATCAACAACCTGATATCCATGAGATATGCGTATATAGTCAGACACGACAC 1429
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 485 CATCAACTACCACTACTACACTACACACTACACTACACTACACTACACTAC 544
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1430 CAACAACAACAACAGTGTGATCGTCCAGTGTCAACAACAACAACAATTAATGT 1489
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 545 CACCAACTACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 604
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1490 GGATCATCCCATGATGATTAACAACAACAACAATTTGACAAACAGGACATACAGAG 1549
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
DB 605 CACCAACAACAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCA 664
    ||| ||| | | | | | | | | | | | | | | | | | | | | |
QY 1550 AGACAA 1555
    ||| |||
DB 665 CGCCAA 670

```

us-09-890-220-1.rtf

RESULT 8
 US-09-216-393-343
 ? Sequence 343 Application US/09216393
 ? Patent No. US20010014447A1
 ? GENERAL INFORMATION
 ? APPLICANT: Milhausen, Michael James
 ? TITLE OF INVENTION: TOXOPLASMA GONDIIT PROTEINS, NUCLEIC ACID MOLECULES, AND
 ? FILE REFERENCE: 7X-1-CC
 ? CURRENT APPLICATION NUMBER: US/09/216,393
 ? EARLIER FILING DATE: 1998-12-18
 ? EARLIER APPLICATION NUMBER: 08/994,825
 ? NUMBER OF FILING DATE: 1997-12-19
 ? SOFTWARE: SEQ ID NOS: 364
 ? SEQ ID NO 343
 ? Length: 1397
 ? TYPE: DNA
 ? ORGANISM: Toxoplasma gondii
 ? FEATURE:
 ? NAME/KEY: CDS
 ? LOCATION: (238)..(1104)
 ? US-09-216-393-343

RESULT 9
 US-09-216-393-345/c
 / Sequence 345 Application US/09216393
 / Patent No. US20060014447A1
 / GENERAL INFORMATION:
 / APPLICANT: Milbarten, Michael James
 / TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS,
 / FILE REFERENCE: 7X-1-02 US\$ THEREOF
 / CURRENT APPLICATION NUMBER: US/09/216,393
 / EARLIER FILING DATE: 1998-12-18
 / EARLIER APPLICATION NUMBER: 08/994,825
 / NUMBER OF SEQ ID NOS: 364
 / SOFTWARE: Patentin Ver. 2.0
 / SEQ ID NO 345
 / LENGTH: 1397
 / TYPE: DNA
 / ORGANISM: Toxoplasma gondii
 US-09-216-393-345

```

Query Match      3.0%; Score 51.6; DB 10; Length 1397;
Best Local Similarity 54.8%; Pled NO 0.0085;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
OY 1370 CATCAACAATGCAATVACCATCTCGAGATTCCTATATACCTACAGACACACACAC 1429
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 10
US-09-938-842A-3152
/ Sequence 3152, Application US/0993842A
/ Patent No. US20010160378a1
/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Krepes, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ TITLE OF INVENTION: SAME, AND METHODS OF USE
/ FILE REFERENCE: SCRIPT300.3
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US-09/938,842A
/ PRIOR FILING DATE: 2000-08-23
/ PRIOR APPLICATION NUMBER: US 60/227,866
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/264,647
/ PRIOR FILING DATE: 2001-06-22
/ NUMBER OF SEQ ID NOS: 5379
/ SEQ ID NO 3152
/ LENGTH: 333
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-09-938-842A-3152

```

Query Match	3.0%;	Score 51.4;	DB 9;	Length 333;
Best Local Similarity	90.2%;	Pred. No. 0.0039;	Mismatches 6;	Gaps 0;
Matches	55;	Conservative 0;		
QY	1003	CTGAAGCACCCTACTTCTTTCGAAAGCCCAATTTCATTCATTTCTACAGAGTCCAGCCCA		
Db	80	CTTCAGCCGACCTACTTCTTTCGAAAGCCCAATTTCATTCATTTCTACAGAGTCCAGGTA		1062
QY	1063	T 1063		
Db	140	T 140		

US-92-801-368-371
 : Sequence 371, Application us/09801368
 : Patent No. US20020128250A1
 GENERAL INFORMATION:
 APPLICANT: Busby, Robert
 APPLICANT: Call, Brian
 APPLICANT: Hecht, Peter
 APPLICANT: Holtzman, Doug
 APPLICANT: Madden, Kevin
 APPLICANT: Maxon, Mary
 APPLICANT: Milne, Todd
 APPLICANT: No. US20020128250A1man, Thea
 APPLICANT: Royer, John
 APPLICANT: Salama, Sofie
 APPLICANT: Sherman, Amlt
 APPLICANT: Silva, Jeff
 APPLICANT: Summers, Eric

1 APPLICANT: Zuker, Charles S.
 2 APPLICANT: Zhang, Yifeng
 3 APPLICANT: The Regents of the University of California
 4 TITLE OF INVENTION: Assays for Taste Receptor Cell Specificity
 5 TITLE OF INVENTION: Ion Channel
 6 FILE REFERENCE: 02307E-114910US
 7 CURRENT APPLICATION NUMBER: US/10/026,188
 8 CURRENT FILING DATE: 2001-12-21
 9 PRIOR APPLICATION NUMBER: US 60/259,379
 10 PRIOR FILING DATE: 2000-12-29
 11 NUMBER OF SEQ ID NOS: 8

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rnpb

Page 6

```

1 SOFTWARE: FastSeq for Windows Version 3.0
2 SEQ ID NO 6
3 LENGTH: 155074
4 TYPE: DNA
5 ORGANISM: Homo sapiens
6 FEATURE:
7 OTHER INFORMATION: human genomic region containing lrpc5
8 OTHER INFORMATION: (Human Chromosome 11p15.5 FAC clone pD9J15f1
9 OTHER INFORMATION: containing KVL2T1 gene)
10 HS-10-026-188-6

```

Query Match	2.68;	Score 44;	DB 9;	Length 155074;
Best Local Similarity	59.78;			
Matches 74;	Conservation	Pred. NO.	14;	

[illegible]

```

RESULT 15
US-09-893-519A-82
/ Sequence 82, Application US/09893519A
/ Publication NO. US2003002743A1
GENERAL INFORMATION:
APPLICANT: ARADYS PHARMACEUTICALS, INC.
APPLICANT: THOMPSON, Craig
APPLICANT: MOORE, Jeffrey
APPLICANT: BURMAN, Ed T.
APPLICANT: BRADLEY, John
APPLICANT: DESTIVA, Thamarra
APPLICANT: HARRIS, Sandra
APPLICANT: KOMARITSKY, Svetlana
APPLICANT: MENDILLO, Marc
APPLICANT: MOORE, Daniel
APPLICANT: MCCOY, Melissa
APPLICANT: SANDERSON, Karen
APPLICANT: HAO, Tariq
APPLICANT: ZHU, Shuhao
APPLICANT: LONG, Fan
APPLICANT: DAVIDOV, Eugene
TITLE OF INVENTION: ANTIPAINL COMPOUNDS AND METHODS OF USE
FILE REFERENCE: 0342/16548-US2
CURRENT APPLICATION NUMBER: US/09/893, 519A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 82
LENGTH: 1560
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Genbank/CA96630.1
DATABASE ENTRY DATE: 1997-08-11
RELEVANT RESIDUES: (1)..(1560)
US-09-893-519A-82

```

Query Match	2.5%	Score 43.4	DB 9	Length 1560
Best Local Similarity	49.8%	Pred. No. 1.1		
Matches 110; Conservative	0	Mismatches 111	Indels 0	Gaps 0

Search completed: June 19, 2003, 23:44:22
Job time : 291 secs

Search completed: June 19, 2003, 23:44:22
Job time : 291 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 20:03:46 ; Search time 4023 Seconds
(Without alignments)
10761.960 Million cell updates/sec

Title: US-09-890-220-1
Perfect score: 1722
Sequence: 1 caagctcttcatttgct.....aatgattctgtataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

```

Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

```

Result No.	Score	Query Match	Length	DB ID	Description
1	1722	100.0	1722	33 US-09-890-220-1	Sequence 1, Appl1
2	1722	100.0	1722	33 US-09-890-475-57	Sequence 57, Appl1
3	1720.4	99.9	1722	33 US-09-890-220-47	Sequence 47, Appl1
4	1569.6	91.1	1715	33 US-09-890-220-4	Sequence 4, Appl1
5	1569.6	91.1	1715	33 US-09-890-475-58	Sequence 58, Appl1
6	1537.6	89.3	1737	33 US-09-890-220-7	Sequence 7, Appl1
7	979.8	56.9	1497	19 US-09-513-996A-69038	Sequence 69038, A
8	412.2	23.9	6338	33 US-09-890-220-3	Sequence 3, Appl1
9	366.8	21.3	1722	42 US-10-219-999-5373	Sequence 5373, Appl1
10	326.4	19.0	606	22 US-09-565-309A-32201	Sequence 32201, A
11	319.4	18.5	5895	33 US-09-890-220-6	Sequence 6, Appl1
12	319.4	18.5	201471	20 US-09-534-859-587	Sequence 587, App
13	319.4	18.5	201471	31 US-09-803-736-587	Sequence 587, App
14	314.2	18.2	319	19 US-09-513-996A-69042	Sequence 69042, A
15	314	18.2	2346	74 US-60-303-460-9	Sequence 9, Appl1
16	306.8	17.8	638	22 US-09-565-309A-52491	Sequence 52491, A
17	306.8	17.8	640	22 US-09-565-309A-32200	Sequence 32200, A
18	306.8	17.8	640	22 US-09-565-309A-61982	Sequence 61982, A
19	306.4	17.8	640	22 US-09-595-329A-945	Sequence 945, App
20	306.4	17.8	640	19 US-09-513-996A-41806	Sequence 41806, A
21	291.4	16.9	2280	17 US-09-339-947A-2	Sequence 2, Appl1

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```

22 284.2 16.5 2249 74 US-60-303-460-1 Sequence 1, Appl1
23 264.8 15.4 2145 74 US-60-303-460-13 Sequence 13, Appl
24 257.4 14.9 1841 42 US-10-219-999-16339 Sequence 16339, A
25 257.4 14.9 1841 75 US-60-312-544-4687 Sequence 4687, Ap
26 257.4 14.9 1856 33 US-09-865-439A-20106 Sequence 20106, A
27 257.4 14.8 2459 74 US-60-303-460-5 Sequence 5, Appl1
28 255 14.7 2276 74 US-60-303-460-7 Sequence 7, Appl1
29 252.4 14.7 2248 17 US-09-338-947A-9 Sequence 9, Appl1
30 247.8 14.4 2353 22 US-09-565-309A-32199 Sequence 32199, A
31 243.6 14.1 2256 74 US-60-303-460-11 Sequence 11, Appl
32 242.8 14.1 677 27 US-09-654-617-26623 Sequence 26623, A
33 242.8 14.1 677 27 US-09-684-016-26623 Sequence 26623, A
34 236.2 13.7 2640 17 US-09-398-237-6 Sequence 6, Appl1
35 234 13.6 303 25 US-09-654-617-163164 Sequence 163164,
36 234 13.6 303 27 US-09-684-016-163164 Sequence 163164,
37 230.8 13.4 558 33 US-09-890-220-13 Sequence 13, Appl
38 221.8 12.9 704 74 US-60-303-460-17 Sequence 17, Appl
39 214 12.4 1176 74 US-60-303-460-3 Sequence 3, Appl1
40 210.8 12.2 600 43 US-10-260-238-2870 Sequence 2870, Ap
41 210.8 12.2 600 61 US-60-170-912-179 Sequence 179, Ap
42 210.8 12.2 990 76 US-60-325-448-2870 Sequence 2870, Ap
43 209.8 12.2 990 25 US-09-654-617-386601 Sequence 386601,
44 209.8 12.2 990 27 US-09-684-016-386601 Sequence 386601,
45 209 12.1 531 33 US-09-874-708A-29223 Sequence 29223, A

```

ALIGNMENTS

RESULT 1
US-09-890-220-1

```

; Sequence 1, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/890,220
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentm Ver. 2.1
; SEQ ID NO 1722
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-890-220-1

```

Query Match 100.0%; Score 1722; DB 33; Length 1722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAGCTCTTCATTTTGGCTGCTCTCTTACACAGCAATCGGTGTTGCACTTT 60
DB 1 CAAGCTCTTCATTTTGGCTGCTCTCTTACACAGCAATCGGTGTTGCACTTT 60
QY 61 CAGGCTCATCCAAACATCTATTAAGCATATGCAAGAGGGGCTTAATGTT 120
DB 61 CAGGCTCATCCAAACATCTATTAAGCATATGCAAGAGGGGCTTAATGTT 120
QY 121 GCATTGAGTTATCGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGATTTG 180
DB 121 GCATTGAGTTATCGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGATTTG 180
QY 181 CACTAATCTATATGAGGCTTGAAGTGTGATTAATTAAGCAAGATGTAGGC 240
DB 181 CACTAATCTATATGAGGCTTGAAGTGTGATTAATTAAGCAAGATGTAGGC 240
QY 241 AGAATGTGCGGGAATCTCACCAGGAGGAAGATTTCACTGATGAGATCTCTTA 300

```

```

DB 241 AGAATGTGCGGGAATCTCACCAGGAGGAAGATTTCACTGATGAGATCTCTTA 300
QY 301 TATATTTGAACCTGTTCCATCTATTAATACATTTTCCACTTGGCTCTAGGAAACCT 360
DB 301 TATATTTGAACCTGTTCCATCTATTAATACATTTTCCACTTGGCTCTAGGAAACCT 360
QY 361 CGTTTCTCCAAAGTCTTGAACACTACAAAATTTGAGCAAGCCAAAGAACTCAAGT 420
DB 361 CGTTTCTCCAAAGTCTTGAACACTACAAAATTTGAGCAAGCCAAAGAACTCAAGT 420
QY 421 CTACTGGAGTGTGATTTCACTATTAAGATTTGAATACATTAACAGAAACCTGAAG 480
DB 421 CTACTGGAGTGTGATTTCACTATTAAGATTTGAATACATTAACAGAAACCTGAAG 480
QY 481 TTAGGAGAGATTTGCTTGTCCATTTTGTCTATGCTATGAGTCAAGGCTCAG 540
DB 481 TTAGGAGAGATTTGCTTGTCCATTTTGTCTATGCTATGAGTCAAGGCTCAG 540
QY 541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTTTGAGAAAT 600
DB 541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTTTGAGAAAT 600
QY 601 ACCAGACATTAATGTTTCTGTAAACTTAATTCCTTCATATTTTGAGAGAAAGCAAGT 660
DB 601 ACCAGACATTAATGTTTCTGTAAACTTAATTCCTTCATATTTTGAGAGAAAGCAAGT 660
QY 661 ATGAGATAAATTTGACCCCTCTCTCTGCTCGAAACCTGTAAAGCAGCAAGAG 720
DB 661 ATGAGATAAATTTGACCCCTCTCTCTGCTCGAAACCTGTAAAGCAGCAAGAG 720
QY 721 GTGGCAGAAATTAACACAGAGAGACTTAAGTATGCTTTTACGTTGATTCACCAAGT 780
DB 721 GTGGCAGAAATTAACACAGAGAGACTTAAGTATGCTTTTACGTTGATTCACCAAGT 780
QY 781 TAACATATGAGCAGAAATGGAATGATACCTCTTAATGAATGAGAAACCGGTTAGAT 840
DB 781 TAACATATGAGCAGAAATGGAATGATACCTCTTAATGAATGAGAAACCGGTTAGAT 840
QY 841 ATCCGAGGCAACAGAGCTGCTGAGCAATTTGAGATGAGCAGCAACATTCACCAAGCA 900
DB 841 ATCCGAGGCAACAGAGCTGCTGAGCAATTTGAGATGAGCAGCAACATTCACCAAGCA 900
QY 901 TAGCCCACTCTTCTCTGAGAGCTGCTGTAAGTTATATGCAAGCAAGCAAGCTGCTC 960
DB 901 TAGCCCACTCTTCTCTGAGAGCTGCTGTAAGTTATATGCAAGCAAGCAAGCTGCTC 960
QY 961 CTGCTACTAAGACAGAAAGTTATCTGCTGAGCGATGAGAGCTAGAGCCACTACTTC 1020
DB 961 CTGCTACTAAGACAGAAAGTTATCTGCTGAGCGATGAGAGCTAGAGCCACTACTTC 1020
QY 1021 TTCAAGAAACGCAATTTCTATCATCTCACAGAGTCCACCAATGGCGCTGAGCAAGTA 1080
DB 1021 TTCAAGAAACGCAATTTCTATCATCTCACAGAGTCCACCAATGGCGCTGAGCAAGTA 1080
QY 1081 TGTCTGACCGGAGTAGAGAGATGAAGTGAAGTGAAGTGGTCAAGTTTGAAGATCGCC 1140
DB 1081 TGTCTGACCGGAGTAGAGAGATGAAGTGAAGTGAAGTGGTCAAGTTTGAAGATCGCC 1140
QY 1141 AGATGCTTGAATGATTTGAGAGTGAATGAATGAATGAATGAATGAATGAATGAATGA 1200
DB 1141 AGATGCTTGAATGATTTGAGAGTGAATGAATGAATGAATGAATGAATGAATGAATGA 1200
QY 1201 ACTGTTTGAAGAAACAAAGGTTATAGCAGATGTCATATCTCTTGGGCAATGTGAAG 1260
DB 1201 ACTGTTTGAAGAAACAAAGGTTATAGCAGATGTCATATCTCTTGGGCAATGTGAAG 1260
QY 1261 CATTTTCAAGATTTTACAGAGAAAGTGTCAACCGTTACTATCATCTTCTGTGTTGA 1320
DB 1261 CATTTTCAAGATTTTACAGAGAAAGTGTCAACCGTTACTATCATCTTCTGTGTTGA 1320
QY 1321 GATTTGTTTGAATTAACATGAGAACATGAGCTGAGATGAGCAGCATTAACAT 1380
DB 1321 GATTTGTTTGAATTAACATGAGAACATGAGCTGAGATGAGCAGCATTAACAT 1380

```



```

Db      1321 GATTGTTTGTGATTAACATGAGAACGACTGTGACTGACACTGACCAACATCAACAACT 1380
OY      1381 GCAATACCATCTCGAGAAATGGCCCTAATAGTCGAGACACCAACCAACAAACAA 1440
Db      1381 GCAATACCATCTCGAGAAATGGCCCTAATAGTCGAGACACCAACCAACAAACAA 1440
OY      1441 ACAGTGTGATCGTCCGAGTGCATCAACACCAACCAACCAATATGATGATCCCA 1500
Db      1441 ACAGTGTGATCGTCCGAGTGCATCAACACCAACCAACCAATATGATGATCCCA 1500
OY      1501 ATGACATTAACACAAAGAACATTTGACACACAGGACATTAACAGACAGACAAAGTAA 1560
Db      1501 ATGACATTAACACAAAGAACATTTGACACACAGGACATTAACAGACAGACAAAGTAA 1560
OY      1561 TTAATATAGGAAATATCCGGCTTTATATACCGATTAACCGATTAACCTATTCCTC 1620
Db      1561 TTAATATAGGAAATATCCGGCTTTATATACCGATTAACCGATTAACCTATTCCTC 1620
OY      1621 TTTCTTAAAAAATGTTAGGAGCAACAAATTTTATATGTTAGTATCAACTGAT 1680
Db      1621 TTTCTTAAAAAATGTTAGGAGCAACAAATTTTATATGTTAGTATCAACTGAT 1680
OY      1681 TACATTTTATGTTAAAAAATGATGATTCCTGTTAACT 1722
Db      1681 TACATTTTATGTTAAAAAATGATGATTCCTGTTAACT 1722

```

RESULT 2

```

US-09-890-475-57
; Sequence 57, Application US/09890475
; GENERAL INFORMATION:
; APPLICANT: Johanson, Urban
; APPLICANT: West, Joanne
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late floweri
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,475
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/GB00/00197
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: GB 9902660.1
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Landsberg
; OTHER INFORMATION: erecta VRN2 CDNA
US-09-890-475-57

```

```

Query Match      100.0%; Score 1722; DB 33; Length 1722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 CAAGCTTCTCAATTTGCTGCTCTCTCTACACAGCAATCGGTGTTTGGCGCTTT 60
Db      1 CAAGCTTCTCAATTTGCTGCTCTCTCTACACAGCAATCGGTGTTTGGCGCTTT 60
OY      61 CAGGCTCAATCCAGACATCTATATAGCATATTGAGAGAGCGGCTTCTATTTGT 120
Db      61 CAGGCTCAATCCAGACATCTATATAGCATATTGAGAGAGCGGCTTCTATTTGT 120
OY      121 GCATTGAGTTATCGCTATGAGAGGAAATTTCAATTTAGGGAGCGCTTCAGAGTTG 180
Db      121 GCATTGAGTTATCGCTATGAGAGGAAATTTCAATTTAGGGAGCGCTTCAGAGTTG 180
OY      181 CACTTAAGTCAATATCGGCTCTTGACGTTGTTGATGTAATTTGAACAAGATGTAGGC 240
Db      181 CACTTAAGTCAATATCGGCTCTTGACGTTGTTGATGTAATTTGAACAAGATGTAGGC 240

```

```

OY      241 AGAATTTGCGCGGAAATCTTCACCGGAGGAAGTATTTCAACTGATGAAATCTCTTGA 300
Db      241 AGAATTTGCGCGGAAATCTTCACCGGAGGAAGTATTTCAACTGATGAAATCTCTTGA 300
OY      301 TATATTTAAACCTGTTGACTATATTAACATCTTTACCTTGCCTCTAGCAACCCAT 360
Db      301 TATATTTAAACCTGTTGACTATATTAACATCTTTACCTTGCCTCTAGCAACCCAT 360
OY      361 CGTTTCTTCCAAAGATGCTTGAACATCAAAATTTGGAGCAAGCGCAAAAGAAAGTCAAGT 420
Db      361 CGTTTCTTCCAAAGATGCTTGAACATCAAAATTTGGAGCAAGCGCAAAAGAAAGTCAAGT 420
OY      421 CTACTGGATGATGTTTTCACATTAAGATTTGATTAACATTAACATTAACAGAAATCTGAAG 480
Db      421 CTACTGGATGATGTTTTCACATTAAGATTTGATTAACATTAACATTAACAGAAATCTGAAG 480
OY      481 TTAGGAGAGATGTTCTTGTCCATTTTCTATGCTATGCTATGCTTCAAGGCGCTGC 540
Db      481 TTAGGAGAGATGTTCTTGTCCATTTTCTATGCTATGCTATGCTTCAAGGCGCTGC 540
OY      541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAGATTTTCCAGAAAT 600
Db      541 AATTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAGATTTTCCAGAAAT 600
OY      601 ACCAGACAGTTAATGTTTCTGTAAACCTTAATCTCATATTTAGAGAAAGAGAGTGG 660
Db      601 ACCAGACAGTTAATGTTTCTGTAAACCTTAATCTCATATTTAGAGAAAGAGAGTGG 660
OY      661 ATGACGATTAATTTGAGCCCTTCTCTCTGCTGCAAAACCTGTAAGCGGAGACAAAGTG 720
Db      661 ATGACGATTAATTTGAGCCCTTCTCTCTGCTGCAAAACCTGTAAGCGGAGACAAAGTG 720
OY      721 GTGGAGAAATTAACACCAAGAGACTTAAGATGTTTACCCTGATTCACCACTT 780
Db      721 GTGGAGAAATTAACACCAAGAGACTTAAGATGTTTACCCTGATTCACCACTT 780
OY      781 TAACTAATGGCAGACAAATGGAATCACCCTACTTAATGATGAAACCGTGGTTAGAT 840
Db      781 TAACTAATGGCAGACAAATGGAATCACCCTACTTAATGATGAAACCGTGGTTAGAT 840
OY      841 ATCCGAGCAACAGAGCTTGTGAGCAATTTGAGATGACAGCAACATTCACAGAGCA 900
Db      841 ATCCGAGCAACAGAGCTTGTGAGCAATTTGAGATGACAGCAACATTCACAGAGCA 900
OY      901 TAGCCCACTCTTCTGTGACGCTGTGCTAAAGTATTAATGACAAAGCAAGCTGTGCTC 960
Db      901 TAGCCCACTCTTCTGTGACGCTGTGCTAAAGTATTAATGACAAAGCAAGCTGTGCTC 960
OY      961 CTGCTACTAAGACAGAAAGTATATCTGCTGAGCATCAGAGGCTGAGAGCAACCTACTTC 1020
Db      961 CTGCTACTAAGACAGAAAGTATATCTGCTGAGCATCAGAGGCTGAGAGCAACCTACTTC 1020
OY      1021 TTCAAGAAAGCCCAATTTCTATATCTCTACAGAGTCCACCAATGCGCTTGGAGCAATGA 1080
Db      1021 TTCAAGAAAGCCCAATTTCTATATCTCTACAGAGTCCACCAATGCGCTTGGAGCAATGA 1080
OY      1081 TGTCTGACCGGGATGAGAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140
Db      1081 TGTCTGACCGGGATGAGAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140
OY      1141 AGATGCTTGATGACTTGTGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1200
Db      1141 AGATGCTTGATGACTTGTGATGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1200
OY      1201 ACTGTTTGTAAAGAAACAAAGGTTATAGCAGATGCTATATCTCTGGGAGATGTAAG 1260
Db      1201 ACTGTTTGTAAAGAAACAAAGGTTATAGCAGATGCTATATCTCTGGGAGATGTAAG 1260
OY      1261 CATTTTCAAGATTTTACGAGAAAGATGACCGTATCATCATCATCTTCTGGTGTGGA 1320
Db      1261 CATTTTCAAGATTTTACGAGAAAGATGACCGTATCATCATCATCTTCTGGTGTGGA 1320
OY      1321 GATGTTTGTGATTAACATATGAGAACATGACCTGTGACTCAGCCACCATCAACAACT 1380

```

```

Db 1321 GATTGTTTGGTAAACATGAAACGATGAGCTGAGCTACAGCCACCCTACAACT 1380
QY 1381 GCAATACCATCTCGAGAAATTCGGTAATAGCTACAGACACCACCACCACAAACAA 1440
Db 1381 GCAATACCATCTCGAGAAATTCGGTAATAGCTACAGACACCACCACCACAAACAA 1440
QY 1441 ACAAGTGTGATCGCTCCAGTCACTCAACACCAACAAATTAACATTTGATGATCCCA 1500
Db 1441 ACAAGTGTGATCGCTCCAGTCACTCAACACCAACAAATTAACATTTGATGATCCCA 1500
QY 1501 ATGCATATTAACCAAGAACATGTTGACACAGACAAATTAACAGACAGACAAAGTAA 1560
Db 1501 ATGCATATTAACCAAGAACATGTTGACACAGACAAATTAACAGACAGACAAAGTAA 1560
QY 1561 TTAATATGAGAAATTCGCGCTTTATGATACGATTTATCGATTTGATATTTCTTC 1620
Db 1561 TTAATATGAGAAATTCGCGCTTTATGATACGATTTATCGATTTGATATTTCTTC 1620
QY 1621 TTCTTAAAAAATGTTTGAAGAGCAAAATTTTATATGTTAGTATCACTGAT 1680
Db 1621 TTCTTAAAAAATGTTTGAAGAGCAAAATTTTATATGTTAGTATCACTGAT 1680
QY 1681 TACATTTTATGTTAAAAAATGATTTCTGTTAACT 1722
Db 1681 TACATTTTATGTTAAAAAATGATTTCTGTTAACT 1722

```

RESULT 3
US-09-890-220-47
Sequence 47, Application US/09890220
GENERAL INFORMATION:

APPLICANT: Dean, Caroline
APPLICANT: Gendall, Anthony
TITLE OF INVENTION: Methods and means for modification of plant characteristics using
FILE REFERENCE: Newbury
CURRENT APPLICATION NUMBER: US/09/890,220
PRIORITY FILING DATE: 2001-07-27
PRIORITY APPLICATION NUMBER: PCT/GB00/00248
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: GB 9901927.5
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 47
LENGTH: 1722
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-47

Query Match 99.9%; Score 1720.4; DB 33; Length 1722;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CAAGCTTCTCAATTTGCTGCTCTCTTACACAGCCCAATGCGCTTTGCGAGCTT 60
Db 1 CAAGCTTCTCAATTTGCTGCTCTCTTACACAGCCCAATGCGCTTTGCGAGCTT 60
QY 61 CAGGCTCAATCCAAAGACATTTCTATATAGCATTTGCGAGAGGCGGTTTAATTTG 120
Db 61 CAGGCTCAATCCAAAGACATTTCTATATAGCATTTGCGAGAGGCGGTTTAATTTG 120
QY 121 GCATTTAGTTTATCGCTTATGACGTAGAGAAATTTCTAATTTAGGGAGGCTCAGAGTTG 180
Db 121 GCATTTAGTTTATCGCTTATGACGTAGAGAAATTTCTAATTTAGGGAGGCTCAGAGTTG 180
QY 181 CACTAATTCATATACGCTCTGACGTTGTTAGTGTATTAACAAAGATGTAGC 240
Db 181 CACTAATTCATATACGCTCTGACGTTGTTAGTGTATTAACAAAGATGTAGC 240
QY 241 AGAATGTGCGCGGAAATCTCACCGGAGAGATTCATGATGATGATCTCTG 300
Db 241 AGAATGTGCGCGGAAATCTCACCGGAGAGATTCATGATGATGATCTCTG 300

```

```

Db 241 AGAATGTGCGCGGAAATCTCACCGGAGAGATTCATGATGATGATGATCTCTG 300
QY 301 TATATTTTAAACGTTTGCATATATATACATTTTACCTTGTCTTATAGGACCAT 360
Db 301 TATATTTTAAACGTTTGCATATATATACATTTTACCTTGTCTTATAGGACCAT 360
QY 361 CGTTCTTCCAAAGATGCTTGAACATGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 CGTTCTTCCAAAGATGCTTGAACATGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 CTACGGGATGATGATTTTCACTATATATATATATATATATATATATATATAT 480
Db 421 CTACGGGATGATGATTTTCACTATATATATATATATATATATATATATATAT 480
QY 481 TTAGGAGAGATGATGATTTTCACTATATATATATATATATATATATATATAT 540
Db 481 TTAGGAGAGATGATGATTTTCACTATATATATATATATATATATATATATAT 540
QY 541 AATTTCAATTTGATTCATCTCATGATTTATTTGAATTTGATGATGATGATGAT 600
Db 541 AATTTCAATTTGATTCATCTCATGATTTATTTGAATTTGATGATGATGATGAT 600
QY 601 ACCAGACAGTTAATGTTTCTGTAATAATTCCTTCATATTTGAGAGAGAGAGAG 660
Db 601 ACCAGACAGTTAATGTTTCTGTAATAATTCCTTCATATTTGAGAGAGAGAGAG 660
QY 661 ATGACAGATTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 ATGACAGATTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GTGCGAGAGATTAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 GTGCGAGAGATTAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 TAATATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 TAATATATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 ATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 ATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 TAGCCACCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 TAGCCACCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CTGCTACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 CTGCTACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 TTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 TGTCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 TGTCTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 AGATGCTTGAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 AGATGCTTGAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ACTGCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 ACTGCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 CATTTTCAAGATTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 CATTTTCAAGATTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 GATGCTTTTGAATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 GATGCTTTTGAATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

```

```

QY 1381 GCAATACATCCCTCGAATGCGCGTAATAGCTCAGACACACACACACCAACCAACACACA 1440
DB 1381 GCAATACATCCCTCGAATGCGCGTAATAGCTCAGACACACACACACCAACCAACCAACA 1440
QY 1441 ACAGTGTGATGCTGCCAGTGTGACTCAACACACCAACAATTAACATTTGGATTCATCCCA 1500
DB 1441 ACAGTGTGATGCTGCCAGTGTGACTCAACACACCAACAATTAACATTTGGATTCATCCCA 1500
QY 1501 ATGACATTAACAACAAGAAACAATGTTGACACAGACACATTAACAGACAGACAAGTAA 1560
DB 1501 ATGACATTAACAACAAGAAACAATGTTGACACAGACACATTAACAGACAGACAAGTAA 1560
QY 1561 TTAATPAGGAATATCGCGCTTTATATGATACCGANTTTGCGATTTGATTAATTTCTTC 1620
DB 1561 TTAATPAGGAATATCGCGCTTTATATGATACCGANTTTGCGATTTGATTAATTTCTTC 1620
QY 1621 TTTCTTAAATAATGTTTAGAGCAACAATTTTAAATGTTAGTGTATTCAACTGAT 1680
DB 1621 TTTCTTAAATAATGTTTAGAGCAACAATTTTAAATGTTAGTGTATTCAACTGAT 1680
QY 1681 TACATTTTATGTTAATAAATAATGATTCGCTTATTAAT 1722
DB 1681 TACATTTTATGTTAATAAATAATGATTCGCTTATTAAT 1722

```

RESULT 4

```

US-09-890-220-4
; Sequence 4, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; APPLICANT: Gendall, Anthony
; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
; FILE REFERENCE: Mewdurn
; CURRENT APPLICATION NUMBER: US/09/890,220
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-890-220-4

```

Query Match 91.1%; Score 1569.6; DB 33; Length 1715;
 Best Local Similarity 96.5%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

```

QY 1 CAAGCTTCTTCAATTTTGGTTCG-CTCTCTCTTACACAGCAATCGTGTGTTTCGAGCT 58
DB 1 CAAGCTTCTTCAATTTTGGTTCGCTCTCTCTTACACGCGCAATCGGTGTTTCGAGCT 60
QY 59 TTCAGGCTTCATCCAGACATTTCTATATACCATATTCAGACAGAGGCGGTTCTAATG 118
DB 59 TTCAGGCTTCATCCAGACATTTCTATATACCATATTCAGACAGAGGCGGTTCTAATG 118
QY 61 TTCAGGCTTCATCCAGACATTTCTATATACCATATTCAGACAGAGGCGGTTCTAATG 120
DB 61 TTCAGGCTTCATCCAGACATTTCTATATACCATATTCAGACAGAGGCGGTTCTAATG 120
QY 119 TTGCAATGATTTATCGTATGACGATGAGGAAATCTATTTAGGGGAGGCGTCAGAGT 178
DB 119 TTGCAATGATTTATCGTATGACGATGAGGAAATCTATTTAGGGGAGGCGTCAGAGT 180
QY 121 TTGCAATGATTTATCGTATGACGATGAGGAAATCTATTTAGGGGAGGCGTCAGAGT 180
DB 121 TTGCAATGATTTATCGTATGACGATGAGGAAATCTATTTAGGGGAGGCGTCAGAGT 180
QY 179 TGCATTAATCTATATGAGGCTCTTGCAGCTTGTGAGTGTATTTGAACAAGATGTGAG 238
DB 179 TGCATTAATCTATATGAGGCTCTTGCAGCTTGTGAGTGTATTTGAACAAGATGTGAG 240
QY 181 TGCATTAATCTATATGAGGCTCTTGCAGCTTGTGAGTGTATTTGAACAAGATGTGAG 240
DB 181 TGCATTAATCTATATGAGGCTCTTGCAGCTTGTGAGTGTATTTGAACAAGATGTGAG 240
QY 239 GCAGAAATGTGCGGGAATCTCTCAACGAGAGAGTGTATTCAACTGATGAGAAATCTCT 298
DB 239 GCAGAAATGTGCGGGAATCTCTCAACGAGAGAGTGTATTCAACTGATGAGAAATCTCT 298
QY 241 GCAGAAATGTGCGGGAATCTCTCAACGAGAGAGTGTATTCAACTGATGAGAAATCTCT 300
DB 241 GCAGAAATGTGCGGGAATCTCTCAACGAGAGAGTGTATTCAACTGATGAGAAATCTCT 300

```

```

QY 299 GATATATGTAAACCTGTTGCACTATATTAACATCTTCACTTCGCTCTGAGGAACCC 358
DB 301 GATATATGTAAACCTGTTGCACTATATTAACATCTTCACTTCGCTCTGAGGAACCC 360
QY 359 ATCGTTCCTCCAGATGCTTGAACATTAACAAATTTGGAGCAAGCGCAAAAGAAAGTCAAG 418
DB 361 ATCGTTCCTCCAGATGCTTGAACATTAACAAATTTGGAGCAAGCGCAAAAGAAAGTCAAG 420
QY 419 ATCTACTGGGATGATGATTTTCACTATATAGATTTGTATTAACATTAACAGAAACTGA 478
DB 421 ATCTACTGGGATGATGATTTTCACTATATAGATTTGTATTAACATTAACAGAAACTGA 480
QY 479 AGTTAGGGAGATGTTGTTCTGCAATTTGCTATGATGATGATGATGATGATGATGATGAT 538
DB 481 AGTTAGGGAGATGTTGTTCTGCAATTTGCTATGATGATGATGATGATGATGATGATGAT 540
QY 539 GCAATTTCAATTTGAATTCATCATGATTTATTTGAATTTGATTTGATTTGATTTGATTT 598
DB 541 GCAATTTCAATTTGAATTCATCATGATTTATTTGAATTTGATTTGATTTGATTTGATTT 600
QY 599 ATACCAAGACATTAATGTTTCTGTAATTAATTTCTCATATTTGAGGAAGAAAGAAAG 658
DB 601 ATACCAAGACATTAATGTTTCTGTAATTAATTTCTCATATTTGAGGAAGAAAGAAAG 660
QY 659 TGATGACATTAATTTGACCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 718
DB 661 TGATGACATTAATTTGACCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 719 AGGTGACAGAAATTAACACACAGAGACATTAAGTATGCTTTTAACTGATTAACACACAG 778
DB 721 AGGTGACAGAAATTAACACACAGAGACATTAAGTATGCTTTTAACTGATTAACACACAG 780
QY 779 TTTAATTAATGACAGAAATTAAGTATGCTTTTAACTGATTAACACACAGAGAGAGAG 838
DB 781 TTTAATTAATGACAGAAATTAAGTATGCTTTTAACTGATTAACACACAGAGAGAGAG 840
QY 839 ATATCCCGAGGCAACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898
DB 841 ATATCCCGAGGCAACAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 899 CATAGCCACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
DB 901 CATAGCCACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 959 CCTGCTACTAGACAGAAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1018
DB 961 CCTGCTACTAGACAGAAAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1019 TCTTCAGAAAGCCCAATTTATCATTTCTACAGAGTCCAGCCCAATGAGGCTTGAGCACT 1078
DB 1021 TCTTCAGAAAGCCCAATTTATCATTTCTACAGAGTCCAGCCCAATGAGGCTTGAGCACT 1080
QY 1079 AATGCTGACCGGATGAGGAGATGAGATGATGATGATGATGATGATGATGATGATGATG 1138
DB 1081 AATGCTGATCGGATGAGGAGATGAGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1139 CCAGATGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1198
DB 1141 CCAGATGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1199 GAATCTGTTTGAAGAAACAAGGTTTATACAGATGCTATCTGTTGGGCAATGTA 1258
DB 1201 GAATCTGTTTGAAGAAACAAGGTTTATACAGATGCTATCTGTTGGGCAATGTA 1260
QY 1259 AGCATTTTCAAGATTTTACAGAAAGTGTGACCGTTATCTCATCTCTTGTGTTG 1318
DB 1261 AGCATTTTCAAGATTTTACAGAAAGTGTGACCGTTATCTCATCTCTTGTGTTG 1320
QY 1319 GAGATTTGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1378
DB 1321 GAGATTTGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
QY 1379 CTGCAATACATCTCTGAGAAATTTGCCGTATAGCTCAGACACACACACCAACCAACA 1438

```

```

Db      1381  CTGCAATTCCTCCGAGAAATTCGCGTAATACCTCAG-----TCACATACAGACAGAA 1434
Oy      1439  CAACAGTGTGATGTCGCCAGTACTCAACACCAACAAACAAATATGATGATGATC 1498
Db      1435  CAACAGTGTGATGTCGCCAGTACTCAACACCAACCAACAAATATGATGATC 1494
Oy      1499  CAATGACATTAACAAACAAACAAATGTTGACACAGACAAATATGACAGACAAAG 1558
Db      1495  GAATGACATTAACAAACAAACAAATGTTGACACAGACAAATATGACAGACAAAG 1553
Oy      1559  AATTAAATAGGAAATCTCCGCTTTATGATACCAATTTATGATGATGATGATCT 1618
Db      1554  AATTAAATAGGAAACCTCCGCTTTATGATACCAATTTATGATGATGATGATCT 1613
Oy      1619  TCTTCTT-AAAAATGTTAGAGCAACAAATTTTATGATGATGATGATGATCACT 1677
Db      1614  TCTTCTTAAAAAATTTTATGAGCAACAAATTTTATGATGATGATGATCACT 1673
Oy      1678  GATTACATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705
Db      1674  GATTACATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1701

RESULT 5
US-09-890-475-58
; Sequence 58, Application US/09890475
; GENERAL INFORMATION:
; APPLICANT: Johnson, Urban
; APPLICANT: West, Joanne
; APPLICANT: West, Joanne
; TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late flower
; FILE REFERENCE: US/09/890,475
; CURRENT FILING DATE: 2001-11-13
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 1/25
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent Ver. 2.1
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Columbia VRN2
US-09-890-475-58

Query Match
Best Local Similarity 91.1%; Score 1569.6; DB 33; Length 1715;
Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

Oy      1  CAAGCTTCTCAATTTGCTTG--CTCTCTTACACAGCCAAATGCTGTTTGCAGCT 58
Db      1  CAAGCTTCTCAATTTGCTTGCTCTCTTACACAGCCAAATGCTGTTTGCAGCT 60
Oy      59  TTCAGGCTCAATCCAAAGACATCTATATTAAGCATATTCGAAGAGCGGCTTCAAT 118
Db      61  TTCAGGCTCAATCCAAAGACATCTATATTAAGCATATTCGAAGAGCGGCTTCAAT 120
Oy      119  TTGCATGAGTTATGCTATGACGTAGAGGAAATTTCAATTTAGGGAGGCTCAGAGT 178
Db      121  TTGCATGAGTTATGCTATGACGTAGAGGAAATTTCAATTTAGGGAGGCTCAGAGT 180
Oy      179  TGCATAACTTCAATTCGGCTGCTGACGTGCTGATGATGATGATGATGATGATG 238
Db      181  TGCATAACTTCAATTCGGCTGCTGACGTGCTGATGATGATGATGATGATGATG 240
Oy      239  GCAGAAATGTCGGCGGAATCTCCACGGAGAGAGTATTTCAATGATGATGATGAT 298
Db      241  GCAGAAATGTCGGCGGAATCTCCACGGAGAGAGTATTTCAATGATGATGATGAT 300

Oy      289  GATATATGTAACCTGTTGACATATTAATCAATCTTACCTGCTGCTATAGCAACCC 358
Db      301  GATATATGTAACCTGTTGACATATTAATCAATCTTACCTGCTGCTATAGCAACCC 360
Oy      359  ATGCTTCTTCCAAAGATGCTGATCAATTAATGATGATGATGATGATGATGATG 418
Db      361  ATGCTTCTTCCAAAGATGCTGATCAATTAATGATGATGATGATGATGATGATG 420
Oy      419  ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 478
Db      421  ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Oy      479  AGTTAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 538
Db      481  AGTTAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Oy      539  GCATTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 598
Db      541  GCATTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Oy      599  ATACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 658
Db      601  ATACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Oy      659  TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 718
Db      661  TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Oy      719  AGTGGGCAAAATTAACCCAGGAGCTTAAGATGATGATGATGATGATGATGATG 778
Db      721  AGTGGGCAAAATTAACCCAGGAGCTTAAGATGATGATGATGATGATGATGATG 780
Oy      779  TTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
Db      781  TTATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Oy      839  ATATCCGAGGCAAGAGCTGCTGAGCAATTTGATGATGATGATGATGATGATGAT 898
Db      841  ATATCCGAGGCAAGAGCTGCTGAGCAATTTGATGATGATGATGATGATGATGAT 900
Oy      899  CATAGCCCACTCTCTGAGAGCGGTGCTTAAGTATTTGATGATGATGATGATGAT 958
Db      901  CATAGCCCACTCTCTGAGAGCGGTGCTTAAGTATTTGATGATGATGATGATGAT 960
Oy      959  CCTGCTACTAGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
Db      961  CCTGCTACTAGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Oy      1019  TCTTCAGAAACCCCAATTTCTATCTCAGAGTCCAGCAATGAGGCTTGGCAAGT 1078
Db      1021  TCTTCAGAAACCCCAATTTCTATCTCAGAGTCCAGCAATGAGGCTTGGCAAGT 1080
Oy      1079  AATGCTGACGGGATGAGAGTATGATGATGATGATGATGATGATGATGATGATG 1138
Db      1081  AATGCTGATGGATGAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1140
Oy      1139  CCAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
Db      1141  CCAGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Oy      1199  GATCTGTTTGAAGAAACAAAGGTTATATGATGATGATGATGATGATGATGATG 1258
Db      1201  GATCTGTTTGAAGAAACAAAGGTTATATGATGATGATGATGATGATGATGATG 1260
Oy      1259  AGCATTTTCAAGATTTTACAGAAAGATGATGATGATGATGATGATGATGATGAT 1318
Db      1261  AGCATTTTCAAGATTTTACAGAAAGATGATGATGATGATGATGATGATGATGAT 1320
Oy      1319  GAGATGTTTGTGATTAACATGATGATGATGATGATGATGATGATGATGATGATG 1378
Db      1321  GAGATGTTTGTGATTAACATGATGATGATGATGATGATGATGATGATGATGATG 1380

```

QY 1379 CTGCAATACATCTCGAGAAATTCGCTAATAGCTACAGACACCACCACCAACAA 1438
 |||||
 Db 1381 CTGCAATACATCTCGAGAAATTCGCTAATAGCTACAGACACCACCACCAACAA 1434
 |||||
 QY 1439 CAACAGTGTGATGCTCCAGTACTCAACACCAACAAATTAATTTGGATATCC 1498
 |||||
 Db 1435 CAACAGTGTGATGCTCCAGTACTCAACACCAACAAATTAATTTGGATATCC 1494
 |||||
 QY 1499 CAATGACATTAACCAAGAAACAAATGTGACACAAAGACATTAACAGACAGAAAGT 1558
 |||||
 Db 1495 GATATACATTAACCAAGAAACAAATGTGACACAAAGACATTAACAGACAGAAAGT 1553
 |||||
 QY 1559 AATTAATAGGAAATCTCCGGCTTTATGATACCGATTTATCGATTTGATATCTCT 1618
 |||||
 Db 1554 AATTAATAGGAAACACCTCCGGTTATGATGATACCGATTTATCGATTTGATATCTCT 1613
 |||||
 QY 1619 TCTTTCTT-AAAAATTTTGAAGACAAACAAATTTTATATGTATGTATTCACCT 1677
 |||||
 Db 1614 TCTTTCTTAAAAAATTTGTTAGAGACAAACAAAGATTTTATTTGTATGTATTCACCT 1673
 |||||
 QY 1678 GATTACATTTTGTATTAATAAATAAT 1705
 |||||
 Db 1674 GATTACATTTTGTATTAATAAATAATGAT 1701
 |||||

RESULT 6

US-09-890-220-7
 ; Sequence 7, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendall, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; PRIOR FILING DATE: 1999-01-28
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: Patent Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 1737
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-7

Query Match 89.3%; Score 1537.6; DB 33; Length 1737;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 49; Indels 32; Gaps 5;

QY 1 CAAGCTTCTCAATTTTCTG--CTCTCTTACACAGCCAAATCGGTTTTCGACCT 58
 |||||
 Db 1 CAAGCTTCTCAATTTTCTGCTCTCTCTTACAGCCAAATCGGTTTTCGACCT 60
 |||||
 QY 59 TTCAGGCTCAATCCAAACATTTCTATATAGCAATTCGAGAAAGCGGTTCTAATG 118
 |||||
 Db 61 TTCAGGCTCAATCCAAACATTTCTATATAGCAATTCGAGAAAGCGGTTCTAATG 120
 |||||
 QY 119 TTGCAATGATTTATCGCTATGACGTAGGAAATTTCTAATTTAGGAGGCTCGAGAT 178
 |||||
 Db 121 TTGCAATGATTTATGACGTATGACGTAGGAAATTTCTAATTTAGGAGGCTCGAGAT 180
 |||||
 QY 179 TGCATTAATCTCAATATCGGCTCTTACGTTGTTGATGTAATTTGACAAAGATGTAG 238
 |||||
 Db 181 TGCATTAATCTCAATATCGGCTCTTACGTTGTTGATGTAATTTGACAAAGATGTAG 240
 |||||
 QY 239 GCAGATTTGCGCGGAAATCTCTACCGGAGAGATGTTTCACTGATGAGAAATCTCT 298
 |||||
 Db 241 GCAGATTTGCGCGGAAATCTCTACCGGAGAGATGTTTCACTGATGAGAAATCTCT 300
 |||||
 QY 299 GATATATTTAAACCTGTTGACATATATACATCTTACCTTCGCTCTAGCAACCC 358
 |||||

Db 301 GATATATTTAAACCTGTTGACATATATACATCTTACCTTCGCTCTAGCAACCC 360
 |||||
 QY 359 ATCGTTTCTCCAGATGCTTTGACTACAAAATTTGAGCAAAAGCCAAAAGATCAG 418
 |||||
 Db 361 ATCGTTTCTCCAGATGCTTTGACTACAAAATTTGAGCAAAAGCCAAAAGATCAG 420
 |||||
 QY 419 ATCTAGTGGATGATTTTCACTATATAGATTTGATATACATTTACAAAACCTA 478
 |||||
 Db 421 ATCTAGTGGATGATTTTCACTATATAGATTTGATATATACATTTACAAAACCTA 480
 |||||
 QY 479 AGTTAGGAGATGATTTTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 531
 |||||
 Db 481 AGTTAGGAGATGATTTTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 540
 |||||
 QY 532 -----AGGGCTGCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 576
 |||||
 Db 541 CAACATTAACAATGAGGCGCTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 600
 |||||
 QY 577 TTGATTTCAAGCTTTTCAAGATTTCAAGATTTCAAGATTTCAAGATTTCAAGAT 636
 |||||
 Db 601 TTGATTTCAAGCTTTTCAAGATTTCAAGATTTCAAGATTTCAAGATTTCAAGAT 660
 |||||
 QY 637 TCATATTTGAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 696
 |||||
 Db 661 TCATATTTGAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 |||||
 QY 697 AACCTGTAGCGGAGACAAAGAGTGGCAGAAATTAACACAGAGACTTAAAGTATGT 756
 |||||
 Db 721 AACCTGTAGCGGAGACAAAGAGTGGCAGAAATTAACACAGAGACTTAAAGTATGT 780
 |||||
 QY 757 TTTTACCGTGTATTCACCACTTTTACATTAAGGACAGAAATTAACACCACTTAA 816
 |||||
 Db 781 TTTTACCGTGTATTCACCACTTTTACATTAAGGACAGAAATTAACACCACTTAA 840
 |||||
 QY 817 ATGATGAAACCGTGTATTTAGATATCCGAGGACAAAGAGCTCTGACATTTGAGA 876
 |||||
 Db 841 ATGATGAAACCGTGTATTTAGATATCCGAGGACAAAGAGCTCTGACATTTGAGA 900
 |||||
 QY 877 TGACCAACATCTTCCACCAAGCATGATGATGATGATGATGATGATGATGATGATG 936
 |||||
 Db 901 TGACCAACATCTTCCACCAAGCATGATGATGATGATGATGATGATGATGATGATG 960
 |||||
 QY 937 TATTTACAAAGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
 |||||
 Db 961 TATTTACAAAGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 |||||
 QY 997 CAGAGCTAGAAAGCCACTACTTCTTCAAGAAAGCCAAATTTCTATTTCTCAGACATCC 1056
 |||||
 Db 1021 CAGAGCTAGAAAGCCACTACTTCTTCAAGAAAGCCAAATTTCTATTTCTCAGACATCC 1080
 |||||
 QY 1057 AGCCATAGGCGCTTGAAGATATGCTGACCGGATAGCGAGATGATGATGATGATG 1116
 |||||
 Db 1081 AGCCATAGGCGCTTGAAGATATGCTGACCGGATAGCGAGATGATGATGATGATG 1140
 |||||
 QY 1117 ATGTGAGATTTTAAAGTGCAGATGCTGATGATGATGATGATGATGATGATGATG 1176
 |||||
 Db 1141 ATGTGAGATTTTAAAGTGCAGATGCTGATGATGATGATGATGATGATGATGATG 1200
 |||||
 QY 1177 AAAAGCAATTCATGATCTTTGGAATCTGTTTGAAGAAACAAAGGTTTATAGAGATG 1236
 |||||
 Db 1201 AAAAGCAATTCATGATCTTTGGAATCTGTTTGAAGAAACAAAGGTTTATAGAGATG 1260
 |||||
 QY 1237 GTCATATCTCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1296
 |||||
 Db 1261 GTCATATCTCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 |||||
 QY 1297 ACTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
 |||||
 Db 1321 ACTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 |||||
 QY 1357 TCGACTCAGCACATCAACAACTCAATACATCTGAGAAATGCGCTAATAGCTAG 1416
 |||||

Db 1381 TCGACTGAGCCACATCAACAGTGCATACATCTCTGAGAAATTCGCCGTATATACCTCAG 1440
 Oy 1417 ACACGACGACACCAACCAACAAACAGTGTGATGCTCCAGTGCATCAACACCAACA 1476
 Db 1441 -----TCACTAACCAACCAACAAAGTGTGATGCTCCAGTGCATCAACACCAACA 1494
 Oy 1477 ACAATACATTTGTGATGATCCATGACATTAACAACAAGCAATGTGTGACAAG 1536
 Db 1495 ACAATACATTTGTGATGATCCGATGACATTAACAACAAGCAATGTGTGACAAG 1554
 Oy 1537 ACAATACAGACGACAAATATTAATAGAAATATCGCGGCTTTATGATACGAT 1596
 Db 1555 ACAATACAGACGACAC -AAGTATTAATAGAAACATCGCGTTTATGATACGAT 1613
 Oy 1597 TTATGGATTTGATTTATTTCTTTCTT-AAAAATTTTATGAGCAAAATTT 1655
 Db 1614 CTATCGATTTGATTTATTTCTTTCTTTAAAAAATTTTATGAGCAAAATTT 1673
 Oy 1656 TTATATGATGATGATCACTGATTTCAATTTTATTAATAAAAAAAT 1705
 Db 1674 TTATTTGATGATGATCACTGATTTCAATTTTATTAATAAAAAAATGAT 1723

RESULT 7
 US-09-513-996A-69038
 ; Sequence 69038, Application US/09513996A
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-709P
 ; CURRENT APPLICATION NUMBER: US/09/513, 996A
 ; CURRENT FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 81028
 ; SEQ ID NO 69038
 ; LENGTH: 1497
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: 1..1497
 ; OTHER INFORMATION: any n or xaa = unknown
 ; FEATURE:
 ; OTHER INFORMATION: Location 1..1497 / Ceres Seq. ID 2185972
 ; US-09-513-996A-69038

Query Match 56.9%; Score 979.8; DB 19; Length 1497;
 Best Local Similarity 89.7%; Pred. No. 1.1e-218;
 Matches 1118; Conservative 0; Mismatches 27; Indels 102; Gaps 2;

Oy 355 ACCCATCGTTCTTCCAAAGATGCTTGAACATCAAAATTTGAGCAAGCGCAAAAGAAAGT 414
 Db 320 ACCCATCGTTCTTCCAAAGATGCTTGAACATCAAAATTTGAGCAAGCGCAAAAGAAAGT 379
 Oy 415 CAAGATCTACTGGGATGGTGTGTTTCAACTATTAAGATTTGATTAACATTAACAGAAA 474
 Db 380 CAAGATCTACTGGGATGGTGTGTTTCAACTATTAAGATTTGATTAACATTAACAGAAA 439
 Oy 475 CTGAGTTAGGAGAGATGCTTCTGTCATATTTGCTCTATGCTATGCTGATGCTTCAAG 534
 Db 440 CTGAGTTAGGAGAGATGCTTCTGTCATATTTGCTCTATGCTGATGCTTCAAG 499
 Oy 535 GCGTCATTTCAATTTGAAATTCATCTCATGATTTATTAATTTGAGTTCAAGCTTTTCG 594
 Db 500 GCGTCATTTCAATTTGAAATTCATCTCATGATTTATTAATTTGAGTTCAAGCTTTTCG 559
 Oy 595 AAGATACACAGACAGTTAATGTTCTGTAAACTTAATCTTCAATTTGAGAGAGAG 654
 Db 560 CTGTAAATCCAAACATAA----- 577
 Oy 655 GAAGTATGACGATAAATTTGAGCCCTTCTCTGCTGGAACCTGTAAGGGAGAC 714
 Db 578 -----TGTTTGCAGCTCGAAACCTGTAAGGGAGAC 610

Oy 715 AAAGAGTGGCAGAAATTAACACAGAGACTTAAGTATGCTTTTACCTGTGATTCAC 774
 Db 611 AAAGAGTGGCAGAAATTAACACAGAGACTTAAGTATGCTTTTACCTGTGATTCAC 670
 Oy 775 CCAGTTTAACTATGACAGAAATTAAGTATGCTTTTAAATGATGATAACCGTGGTT 834
 Db 671 CCAGTTTAACTATGACAGAAATTAAGTATGCTTTTAAATGATGATAACCGTGGTT 730
 Oy 835 TAGATATCCCGAGGACAGAGAGCTTGTGAGCAATTTGATGATGACAGCAATTCAC 894
 Db 731 TAGATATCCCGAGGACAGAGAGCTTGTGAGCAATTTGATGATGACAGCAATTCAC 790
 Oy 895 CAGCATAGGCGACCTCTCTGAGAGCGTGTGAAAGTATATGACAAGGAAAGCTG 954
 Db 791 CAGCATAGGCGACCTCTCTGAGAGCGTGTGAAAGTATATGACAAGGAAAGCTG 850
 Oy 955 TGGTCCCTGCTACTAAGACAGAAAGTTATCTGCTGAGCGATCAGAGCTAGA----- 1007
 Db 851 TGGTCCCTGCTACTAAGACAGAAAGTTATCTGCTGAGCGATCAGAGCTAGA----- 1007
 Oy 1008 -----AGCCACTACTTCTTTCAGAAACGCCAATTTCTATC 1041
 Db 911 CCATATTTACCGAGAAATTTCTTCCAGCACTACTTCTTTCAGAAACGCCAATTTCTATC 970
 Oy 1042 ATTCTCAGAGTCCAGCCAAATGCGCTTGTGAGCAAGTATGCTGACCGGATAGGAG 1101
 Db 971 ATTCTCAGAGTCCAGCCAAATGCGCTTGTGAGCAAGTATGCTGACCGGATAGGAG 1030
 Oy 1102 ATGAAGTCGATGACGATGTTGAGATTTGAAGATCGCGAGATGCTGATGATTTG 1161
 Db 1031 ATGAAGTCGATGACGATGTTGAGATTTGAAGATCGCGAGATGCTGATGATTTG 1090
 Oy 1162 ATGATTAAGATGAAAGCAATTCATGATGATGATGATGATGATGATGATGATGAT 1221
 Db 1091 ATGATTAAGATGAAAGCAATTCATGATGATGATGATGATGATGATGATGATGAT 1150
 Oy 1222 GGGTTATGACAGATGCTATCTCTGCGCATGGAAGCAATTTTCAAGTTTACGAGA 1281
 Db 1151 GGGTTATGACAGATGCTATCTCTGCGCATGGAAGCAATTTTCAAGTTTACGAGA 1210
 Oy 1282 AAGAGTTGACGCTTACTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1341
 Db 1211 AAGAGTTGACGCTTACTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
 Oy 1342 GGAACATGACCTTGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1401
 Db 1271 GGAACATGACCTTGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 1330
 Oy 1402 GCGTATATGCTGACAGACACACACACACACACACACACACACACACACACACAC 1461
 Db 1331 GCGTATATGCTGACAGACACACACACACACACACACACACACACACACACACAC 1390
 Oy 1462 ACTCAAC 1521
 Db 1391 ACTCAAC 1450
 Oy 1522 ATGTTGACAAAGAGCAATTAACGACAGACAAAGTATTAATTTG 1568
 Db 1451 ATGTTGACAAAGAGCAATTAACGACAGACAAAGTATTAATTTG 1497

RESULT 8
 US-09-890-220-3
 ; Sequence 3, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Gendall, Anthony
 ; APPLICANT: Dean, Caroline
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; FILE REFERENCE: Newbury
 ; CURRENT APPLICATION NUMBER: US/09/890, 220
 ; CURRENT FILING DATE: 2001-07-27

PRIOR APPLICATION NUMBER: PCT/GB00/00248
 PRIOR FILING DATE: 2000-01-28
 PRIOR APPLICATION NUMBER: GB 9901927.5
 PRIOR FILING DATE: 1999-01-28
 NUMBER OF SEQ ID NOS: 77
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 3
 LENGTH: 6338
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-890-220-3

Query Match 23.9%; Score 412.2; DB 33; Length 6338;

Best Local Similarity 98.8%; Pred. No. 2,1e-85;

Matches 414; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1304 ACTCTTCTGCTGTTGAGATTTGTTTGAATTAACACTGGAACCATGACCTTGCACCTC 1363
 DB 5102 AATTTCCAGCTGTGGAGATTTGTTTGAATTAACACTGGAACCATGACCTTGCACCTC 5161
 QY 1364 ACCCAACATCAACACTGCAATACATCTCTCGAATTTGCCGTAATAGCTGACACACCAC 1423
 DB 5162 ACCCAACATCAACACTGCAATACATCTCTCGAATTTGCCGTAATAGCTGACACACCAC 5221
 QY 1424 CACCAACCAACCAACAGTGTGATCGTCCAGTGCATCAACACCAACCAACATATA 1483
 DB 5222 CACCAACCAACCAACAGTGTGATCGTCCAGTGCATCAACACCAACCAACATATA 5281
 QY 1484 CATTTGATCTCCCAATGACATTAACCAACCAACCAATGTTGACACACAGACATATA 1543
 DB 5282 CATTTGATCTCCCAATGACATTAACCAACCAACCAATGTTGACACACAGACATATA 5341
 QY 1544 CAGCAGAGCAAAAGTAATTAATAGAAATCTCGGCTTTATGATCCGATTTATCGG 1603
 DB 5342 CAGCAGAGCAAAAGTAATTAATAGAAATCTCGGCTTTATGATCCGATTTATCGG 5401
 QY 1604 AATGTAATCTTCTCTCTTTTAAATAATGTTAGAGCAACCAATTTTATATGT 1663
 DB 5402 AATGTAATCTTCTCTCTTTTAAATAATGTTAGAGCAACCAATTTTATATGT 5461
 QY 1664 TAGTGTATCACTGATTAATTTTATGTTAAAAAATGATGATTCGCTTATTAAT 1722
 DB 5462 TAGTGTATCACTGATTAATTTTATGTTAAAAAATGATGATTCGCTTATTAAT 5520

RESULT 9

US-10-219-999-5373

Sequence 5373, Application US/10219999

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Edgerton, Michael D

APPLICANT: Hinkle, Gregory J.

APPLICANT: Kovalic, David K.

APPLICANT: Liu, Jingdong

APPLICANT: Stein, Joshua

TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT

FILE REFERENCE: 38-10(52726)C

CURRENT APPLICATION NUMBER: US/10/219, 999

PRIOR FILING DATE: 2002-08-15

PRIOR APPLICATION NUMBER: US 60/324,109

PRIOR FILING DATE: 2001-09-21

PRIOR APPLICATION NUMBER: US 60/312,544

PRIOR FILING DATE: 2001-08-15

NUMBER OF SEQ ID NOS: 63520

SEQ ID NO 5373

LENGTH: 1722

TYPE: DNA

ORGANISM: glycine max

FEATURE:

NAME/KEY: CDS

LOCATION: (164)..(1450)

OTHER INFORMATION:

US-10-219-999-5373

Query Match 21.3%; Score 366.8; DB 42; Length 1722;
 Best Local Similarity 59.8%; Pred. No. 5.7e-75;
 Matches 757; Conservative 0; Mismatches 422; Indels 87; Gaps 5;

QY 228 AGAATGTAGCAGAAATTTGCGCGGAATCTCCACCGAGAGATGATTAACACTGAT 287
 DB 161 AGCATGTGCGGCAAAATTTCTCCGATACACATCCCGGTGAAGAAATAATTTGACATGAT 220
 QY 288 GAGATCTGTGATATATTTGAAACCTGTTCGACTATATATACATCTTTGACCTGT 347
 DB 221 GAGAGCTTTATTTATTTGCAAGCTGTGAACTTTACAAATTTCTCAACGCCCTGCT 280
 QY 348 CTAAGCAACCCATCTTTCTTCCAGATGCTTGAATTAACCAATTTGAGCAAAAGCCAAA 407
 DB 281 CTTCAAAATCTCTTTCTTTGAGATTTGCGGTTATTAATAAGACCAAGCCGTTAA 340
 QY 408 AGAAAGTCAAGATCTACTGGATGTGATTTTCACTATATAGATTTGATTAACATTA 467
 DB 341 AGGAGGTGAG---GCAGGAATTTGATTTCAATTAATAGGACCGTACCAACATTTCT 397
 QY 468 CAGAAATCTGAAGTTAGGAGATTTGTTGCTATTTGCTATATGCTATATGCTATG 527
 DB 398 CGAAAGACTGAAGTGACCAAGACTTTCTGTCTTTGCTTGTATGAGAGTGAGGAAC 457
 QY 528 TTCAGGGGCTGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 587
 DB 458 TTTAAGGTTTGGATTTATCTTTGTTATCTATGATTTGATTTGATTTGATTTGATTT 517
 QY 588 CTTTCGAAGATATACAGACAGTAAATGTTCTGTAATACTTAATCTTCAATTTGAG 647
 DB 518 GTTACTGAAGATTTACCAAGAGTAAATGATTTGATTTGATTTGATTTGATTTGAT 577
 QY 648 GAAAGAGAGTATGATACAT---AAATTTGAGCCCTTCTCTCTGCTGCAACCTGCT 704
 DB 578 AATGTTGCTGATGAGATTAATTCACAAATGCAAACTTTCTTCTGTTCAAGACCTGCA 637
 QY 705 AAGCGGAGCAAAAGAGTGGAGCAAAATTAACAGAGAGACTTAATGATTTGATTTGAG 764
 DB 638 AAACGTGAGAAAGAGACTCTGTTCAATTTGAAAAACGACCAATTAATTTCTGAGAG 697
 QY 765 TTGATTTCAACCA----- 777
 DB 698 TTGACTACCAAGAGCAATATATGATTTCTCAAAAGATGATATATCTATACC 757
 QY 778 -----GTTAATCTATATGACACGAAATTTGAAATACCTTACTATATGA 820
 DB 758 TCCAAGAGAGAAATGTTGTTAGATCTGTTAGAGAAATTTTCCCTAGTGAAGA 817
 QY 821 T-----GGAACCGTGTGTTAGATATCCGAGGCAAGAGCTTCTGACAAATTT 872
 DB 818 AATGACGAGGAAATTTGTTGCTGCTGATCTCTGAGCAATGACCACTGAGAGCATGAG 877
 QY 873 GAGATGACCAAGCAATTCACCA-----GCCATGACCACTTCTCTGACGCT 923
 DB 878 GATTCAGTTTCAACATTCAGAGTGTTCATATGATTCATCCCAATCTCTGAGACCT 937
 QY 924 GTGCTAAAGTTATATGACAAAGCAAGCTGTGCTCTGCTACTATGACAAAGAGTTA 983
 DB 938 GAATGATGATTAATCATATATATTAAGTATCTGCTCTGCTCTGCTCTGCTCTGCTCT 997
 QY 984 TCTGCTGAGCAGTACAGAGCTGAGAGCCACTACTCTTCAAGAAAGCCCAATCTATAT 1043
 DB 998 AGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1057
 QY 1044 TCTCAGAGTCCAGGCAATTTGCGCTTGAAGCAATATATGATGATGATGATGATGAT 1103
 DB 1058 TCACACAGAGTCCAGGCAATTTGCGCTTGAAGCAATATATGATGATGATGATGATGAT 1117
 QY 1104 GAAGTCGATGAGATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGATTT 1163
 DB 1118 GAAGTCGATGAGATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGATTT 1177

```

QY 1164 GTGATTAAGATGAAAGCAATTCATGACATCTTGGAACTCGTTTGAAGAAACAAAGG 1223
DB 1178 GTTCCAAAAGATGAAAGCAATTCATGACATCTTGGAACTCGTTTGAAGAAACAAAGG 1237
QY 1224 GTTATAGCAGATGTCATATCTCTTGGGCAATGTAACATTTTCAAGATTTTACGAGAA 1283
DB 1238 GTGCTAGCAGATGTCATATCTCTTGGGCAATGTAACATTTTCAAGATTTTACGAGAA 1297
QY 1284 GAGTTCACCCGTTTACTCATCTCTTGGTGTGGAGATGTTTGAATTAACATATG 1343
DB 1298 GAGCTGATCTCATCTCTTGGGCAATGTAACATTTTCAAGATTTTACGAGAA 1357
QY 1344 AACCATGAGCTTGTGACTCAGCCACCATCAACCACTGCAATCTCGAATTCG 1403
DB 1358 AATCATGCTCTTGTGAGTCCGTTTGAAGTGAAGTGAATTAACATTTTACGAGAA 1417
QY 1404 CGTAAT 1409
DB 1418 AGAAAT 1423

```

```

RESULT 10
US-09-565-309A-32201
Sequence 32201, Application US/09565309A
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
TITLE OF INVENTION: THEREBY
FILE REFERENCE: 2750-0853P
CURRENT APPLICATION NUMBER: US/09/565,309A
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 68449
SEQ ID NO 32201
LENGTH: 606
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(606)
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
NAME/KEY: misc-feature
LOCATION: (1)..(606)
OTHER INFORMATION: 39533:110036 (Clone Number: Unique Sequence Identifier)
US-09-565-309A-32201

```

```

Query Match 19.0%; Score 326.4; DB 22; Length 606;
Best Local Similarity 98.0%; Pred. No. 1.1e-65;
Matches 341; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1373 CAACAAGTCAATGACATCTGAGAAATTCGCTTAATAGCTCAGACACGACCAACCA 1432
DB 259 CAACAAGTCAATGACATCTGAGAAATTCGCTTAATAGCTCAGACACGACCAACCA 318
QY 1433 CAACAAGTCAATGACATCTGAGAAATTCGCTTAATAGCTCAGACACGACCAACCA 1492
DB 319 CAACAAGTCAATGACATCTGAGAAATTCGCTTAATAGCTCAGACACGACCAACCA 378
QY 1493 TCATCCCAATGACATTAACCAAGAAATGTTGACACAGACATTAACAGCAGAG 1551
DB 379 TCATCCCAATGACATTAACCAAGAAATGTTGACACAGACATTAACAGCAGAG 438
QY 1552 ACAAGTAATTAATAGGAAATTCGCGCTTTATGATACGATTAATCGATTTAATC 1611
DB 439 ACAAGTAATTAATAGGAAATTCGCGCTTTATGATACGATTAATCGATTTAATC 498
QY 1612 TTAATCTCTCTTCTTAAATAATTTTGAAGCAACAAATTTTATATGTTAGTAT 1671
DB 499 TTAATCTCTCTTCTTAAATAATTTTGAAGCAACAAATTTTATATGTTAGTAT 558
QY 1672 TCAACTGATTAATTTTATGTTAGTAAATTTTATGTTAGTAAATTTTATGTTAGT 1719
DB 559 TCAACTGATTAATTTTATGTTAGTAAATTTTATGTTAGTAAATTTTATGTTAGT 606

```

```

RESULT 11
US-09-890-220-6
Sequence 6, Application US/09890220
GENERAL INFORMATION:
APPLICANT: Dean, Caroline
TITLE OF INVENTION: Method and means for modification of plant characteristics us
TITLE OF INVENTION: vernalisation gene VRN2.
FILE REFERENCE: Mewburn
CURRENT APPLICATION NUMBER: US/09/890,220
CURRENT FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
PRIOR FILING DATE: 1999-01-28
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 5895
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-6

```

```

Query Match 18.5%; Score 319.4; DB 33; Length 5895;
Best Local Similarity 92.7%; Pred. No. 1.2e-63;
Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;
QY 1308 TTCTGCTTGGAGATGTTTGTGATTAACATGGAACCATGAGCTTGTGACACGAC 1367
DB 4677 TCCAGGTGTTGAGATGTTTGTGATTAACATGGAACCATGAGCTTGTGACACGAC 4736
QY 1368 ACCATCAACATGCAATACATCTCTCGAATTCGCTTAATAGCTCAGACACACAC 1427
DB 4737 ACCATCAACATGCAATACATCTCTCGAATTCGCTTAATAGCTCAGACACACAC 4790
QY 1428 ACCATCAACATGCAATACATCTCTCGAATTCGCTTAATAGCTCAGACACACAC 1487
DB 4791 ACCATCAACATGCAATACATCTCTCGAATTCGCTTAATAGCTCAGACACACAC 4850
QY 1488 GTGATCATCCCAATGACATTAACCAAGAAATGTTGACAAACGAAATTAACAC 1547
DB 4851 GTGATCATCCCAATGACATTAACCAAGAAATGTTGACAAACGAAATTAACAC 4910
QY 1548 AGAGACAAATTAATAGGAAATTCGCGCTTTATGATACGATTAATCGATTTG 1607
DB 4911 AGAGACAAATTAATAGGAAATTCGCGCTTTATGATACGATTAATCGATTTG 4969
QY 1608 TAACTATTTCTCTTCTT-AAAAATTTGTTAGGACCAACAAATTTTATATGTTAG 1666
DB 4970 TAACTATTTCTCTTCTT-AAAAATTTGTTAGGACCAACAAATTTTATATGTTAG 5029
QY 1667 TGTATTCACGATTAATGATTTTATGTTTAAATTTTAAATTTTAAATTTTAAAT 1705
DB 5030 TGTATTCACGATTAATGATTTTATGTTTAAATTTTAAATTTTAAATTTTAAAT 5068

```

```

RESULT 12
US-09-534-859-587
Sequence 587, Application US/09534859
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Last, Robert L.
APPLICANT: Levin, Irene M.
APPLICANT: Norris, Susan R.
APPLICANT: Parnell, Laurence D.
APPLICANT: Rounsley, Steven D.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
FILE REFERENCE: 38-10(15493)B
CURRENT APPLICATION NUMBER: US/09/534,859
CURRENT FILING DATE: 2000-03-29

```



```

; NUMBER OF SEQ ID NOS: 1127
; SEQ ID NO 587
; LENGTH: 201471
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-534-859-587

```

Query Match	18.5%;	Score 319.4;	DB 20;	Length 201471;
Best Local Similarity	92.7%;	Pred. NO. 4.6e-63;		
Matches 370; Conservative	0;	Mismatches 21;	Indels 8;	Gaps 3;

QY	1308	TTCTGGGTTGGGAGATTTGGTTTTGATTAACATATGAACCATGGAGTTGTGCATCAGCC	1367
Db	42342	TCCAAGTGTGGAGATTTGTTTTTGATTAACCTATGGAACCATGGAGTTGTGCATCAGCC	42401
QY	1368	ACCATCAACATCGCAATACCATCTCTCGAATATTGCGTAAATAGCTCAACACCAACACC	1427
Db	42402	ACCATCAACATCGCAATACCATCTCTCGAATATTGCGTAAATAGCTCAACACCAACACC	42455
QY	1428	ACCAACAACAACAACAGTGTGGATGCTCCAGTACTCAACACCAACAACATAACATT	1487
Db	42456	AACAAACAACAACAACAGTGTGGATGCTCCAGTACTCAACACCAACAACATAACATT	42515
QY	1488	GGGATCATCCCAATGACATTAACAACAAGAACAATGTGACACAAGAACAATAACAC	1547
Db	42516	GGGATCATCCCAATGACATTAACAACAAGAACAATGTGACACAAGAACAATAACAC	42575
QY	1548	AGAGACAAAGTAATTAATAGCAAAATCTCCGGCTTTATGATACCGATTTATCGGATGG	1607
Db	42576	AGAGAC-AACTAATTTAAATAGCAAAACATCTCCGGTTTATGATACCGATCTATCGGATGG	42634
QY	1608	TAACTTATCTCTCTTCTT-AAAAAATGTTTAGAGCAACAATTTTATATGTAG	1666
Db	42635	TAACTTATCTCTCTTCTTAAAAAATGTTTAGAGCAACAACAATTTTATATGTAG	42694
QY	1667	TGATTTCACTGATTTACATTTTATAGTTAAAAAAAAT	1705
Db	42695	TGATTTCACTGATTTACATTTTATAGTTAAAAAAAATGAT	42733

RESULT 13
 US-09-803-736-587
 : Sequence 587, Application US/09803736
 : GENERAL INFORMATION:
 : APPLICANT: Bush, David F.
 : APPLICANT: Levin, Irena M.
 : APPLICANT: Norris, Susan R.
 : APPLICANT: Rounsley, Steven D.
 : APPLICANT: Wiegand, Roger C.
 : TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
 : FILE REFERENCE: 38-10(15493)D
 : CURRENT APPLICATION NUMBER: US/09/803, 736
 : CURRENT FILING DATE: 2001-03-12
 : PRIOR APPLICATION NUMBER: US 09/534, 859
 : PRIOR FILING DATE: 2000-03-29
 : PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
 : PRIOR FILING DATE: 2000-10-20
 : NUMBER OF SEQ ID NOS: 1582
 : SEQ ID NO 587
 : LENGTH: 201471
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 : US-09-803-736-587

	Query Match	Best Local Similarity	18.5%; 92.7%;	Score 319.4; DB 31.	Length 201471;
	Matches	Conservative	0;	Mismatches 21;	Indels 8; Gaps 3;
QY	1308	TTTCGCTGTTGGAGATTGTTTTGATTTAAACTATGCAACCATGGACTTGTGCACCTGACCC			1367
Db	42342	TTCAGGCTGGGAGATTGTTTTGATTTAAACTATGCAACCATGGACTTGTGCACCTGACCC			42401
YY	1368	ACCATCAACAACATGCAATACCATCTCGAGAAATTGGCGTAAATGCTCAGACACACACACC			1427

[illegible]

```

RESULT 14
US-09-513-996A-69042
Sequence 69042, Application US/09513996A
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
TITLE OF INVENTION: ENCODED THEREBY
FILE REFERENCE: 2750-709P
CURRENT APPLICATION NUMBER: US/09/513,996A
CURRENT FILING DATE: 2000-02-25
NUMBER OF SEQ.ID NOS: 81028
SEQ ID NO 69042
LENGTH: 319
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: UNSURE
LOCATION: 1..319
OTHER INFORMATION: any n or xaa = unknown
FEATURE:
OTHER INFORMATION: Location 1..319 / Ceres Seq. ID 2185983
US-09-513-996A-69042

```

Query Match	18.2%	Score 314.2	DB 19	Length 319
Best Local Similarity	99.1%	Pred. No. 6.3e-63		
Matches 316	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY	1375	ACAACGTGAAATACCATCTCGAGAATTCGCGTAATATAGTCTGAGACACACACACACACA	1434	
Db	1	ACAACGTGAAATACCATCTCGAGAATTCGCGTAATATAGTCTGAGACACACACACACACA	60	
QY	1435	ACAACAACAGTGGATGTCGCCAGTACTCAACACCAACAACATAATGATGGATC	1499	
Db	61	ACAACAACAGTGGATGTCGCCAGTACTCAACACCAACAACATAATGATGGATC	120	
QY	1495	ATCCCAATGACATTAACCAACAAGAACATGTTGACACAAAGACACATTAACACAGACAGA	1554	
Db	121	ATCCCAATGACATTAACCAACAAGAACATGTTGACACAAAGACACATTAACACAGACAGA	180	
QY	1555	AAGTAATTAATATAGGAAAATCTCCGGCTTTATGATACCGATTTATCGATTGTAACTTA	1611	
Db	181	AAGTAATTAATATAGGAAAATCTCCGGCTTTATGATACCGATTTATCGATTGTAACTTA	240	
QY	1615	TTCTCTCTTCTTAATAAATGTGTGAGAGCAACAAATTTTATATGTTAGTGATATCA	1674	
Db	241	TTCTCTCTTCTTAATAAATGTGTGAGAGCAACAAATTTTATATGTTAGTGATATCA	300	
QY	1675	ACGTATACATTTTATAGTT	1693	

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rnp

Page 12

Db 301 ACTGATTCATTTCAGTT 319

RESULT 15

US-60-303-460-9

Sequence 9, Application US/60303460

GENERAL INFORMATION:

APPLICANT: Allen, Steve

APPLICANT: Butler, Karlene

APPLICANT: Niu, Xiaping

TITLE OF INVENTION: Flowering Regulating Polynucleotides and Polypeptides

FILE REFERENCE: B81493 US PV

CURRENT APPLICATION NUMBER: US/60/303,460

CURRENT FILING DATE: 2001-07-06

NUMBER OF SEQ ID NOS: 19

SOFTWARE: Microsoft Office 97

SEQ ID NO 9

LENGTH: 2346

TYPE: DNA

ORGANISM: Helianthus sp.

US-60-303-460-9

Query Match 18.28; Score 314; DB 74; Length 2346;

Best Local 58.48; P-adj. 1.5e-62;

Matches 598; Conservative 0; Mismatches 405; Indels 21; Gaps 2;

OY 384 TACAAATTGGAGCAAGCCAAAGAAAGTCAAGATCTACTGGATGGTACTGTTTCAC 443

Db 944 TACAAAGATGCTCTAGACCCGGTGAATGAGATTGCGGTGGAATGTCGTTTAT 1003

OY 444 TATAAGATGTAATAACATACATACAGAAAAGTGAAGTGAAGATGTTCTGTGCA 503

Db 1004 TACAAGTACTACAAATACATACATACAGAGTGAAGATGTTCTGTGCG 1063

OY 504 TTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 563

Db 1064 TTTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 1123

OY 564 GATTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 623

Db 1124 GATCTCTTATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1183

OY 624 AAATTAATTCCTCATATTTGAGAG--AAGGAAGTATGATTAATTTGAGCC 680

Db 1184 AGAAGCATGCAAGCAGCTCTGAGATTAATTCCTGAGACGTTGATCCGAACAATG 1243

OY 681 TTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 740

Db 1244 TTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303

OY 741 AGACTTAAGTATGCTTTTACGCTGATTCACCCAGTTTAACTAATGCGACAGAAAT 800

Db 1304 AATCAT-----GTGATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1345

OY 801 GGAATGACCTACTATATATGTAAGAAACCGTGTATGATATCCGAGCAAGAGCTT 860

Db 1346 GCTCTCAGTACGCTATATGATTAATGATGATGATGATGATGATGATGATGATG 1405

OY 861 GCTGACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 920

Db 1406 TGTAGTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1465

OY 921 GCTGATGCTAAAGTTATATGACAAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTG 980

Db 1466 TTACCTGAAAGCAGCAACCTTGACCTCCAGCCTGCTGCTGCTGCTGCTGCTGCTG 1525

OY 981 TTATCTGCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1040

Db 1526 TTATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1585

OY 1041 CATCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1100

Db 1586 CACTCCACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1645

OY 1101 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1160

Db 1646 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705

OY 1161 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220

Db 1706 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765

OY 1221 AAGGATTAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1280

Db 1766 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1825

OY 1281 AAGATGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340

Db 1826 CAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1885

OY 1341 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1400

Db 1886 TGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1945

OY 1401 TGCC 1404

Db 1946 TACC 1949

Search completed: June 19, 2003, 23:17:45

Job time : 4034 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 19:25:25 ; Search time 2561 Seconds
(without alignments)
10889.742 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctcttcatttgc.....aatgattctgtataact 1722

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	628.2	36.5	672	10	AV822548
2	266.6	15.5	526	17	CNS00MYB
3	237.8	13.8	680	10	AM038171
4	232.4	13.5	769	14	BO505017
5	226	13.1	1128	11	AY104964
6	212	12.3	846	12	BG648271

7	206.2	12.0	480	9	AJ469021
8	203	11.8	569	13	BO611550
9	201.4	11.7	600	13	BI479743
10	198.4	11.5	474	10	AM234600
11	197.8	11.5	391	10	AV817525
12	190	11.0	561	14	BO740672
13	188.8	11.0	467	12	BG580495
14	176.4	10.2	364	9	BI164598
15	174.6	10.1	405	12	BG314044
16	173.6	10.1	554	13	BI784785
17	171.2	9.9	546	13	BI321750
18	170.2	9.9	444	12	BG551103
19	169.6	9.8	612	10	BE203328
20	164	9.5	429	14	BO459233
21	160.6	9.3	445	10	AM202144
22	154.6	9.0	326	9	AI163743
23	144.2	8.4	561	14	BO162202
24	143	8.3	662	14	BO279630
25	132.8	7.7	519	14	BO168835
26	126	7.3	644	13	BI453006
27	125.6	7.3	322	10	AM308408
28	124	7.2	757	13	BO460553
29	121	7.0	515	12	BE481866
30	120.6	7.0	236	14	BO904049
31	113.6	6.6	560	9	AI795561
32	113.6	6.6	590	12	BG649311
33	113.6	6.6	594	10	AM000259
34	110.6	6.4	586	14	BO243123
35	109.2	6.3	727	14	BO801572
36	108.6	6.3	732	14	BO841916
37	105.8	6.1	142	10	AV783345
38	104.8	6.1	463	14	C72616
39	101.2	5.9	576	10	AM076157
40	100.4	5.8	675	13	BI932726
41	94.8	5.5	523	13	BI92698
42	94	5.5	553	12	BG909756
43	94	5.5	553	14	BO620168
44	94	5.5	593	13	BM377912
45	93.6	5.4	540	10	BE587348

ALIGNMENTS

RESULT 1

LOCUS

AV822548 AV822548 RAFUS Arabidopsis thaliana cDNA clone RAF105-08-K19 5', mRNA sequence.

ACCESSION

AV822548 AV822548.1 GI:19864597

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 672)

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shingawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.

REFERENCE

AUTHORS

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: mseki@rcc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a

modified bluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
source Location/Qualifiers

1. 672
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone_id="RAF105-08-K19"
/dev_stage="rosette plants"
/lab_host="SOLR"
/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"

BASE COUNT 197 a 116 c 143 g 216 t

ORIGIN

Query Match 36.5%; Score 628.2; DB 10; Length 672;

Best Local Similarity 97.0%; Pred. No. 3.8e-129; Matches 651; Conservative 0; Mismatches 18; Indels 2; Gaps 1;

2 AAGCTTCTCAATTTTGGTTG--CTCTCTTACACAGCCAAATCGTGTTCGACCTT 59
|||||
2 AAGCTTCTCAATTTTGGTTGCTCTCTCTTACACAGCCAAATCGTGTTCGACCTT 61
DB
60 TCAGGCTTCATCCACAGACATCTATATACCAATTTGCAGAGAGCGGTTCTAATGT 119
DB
62 TCAGGCTTCATCCACAGACATCTATATACCAATTTGCAGAGAGCGGTTCTAATGT 121
DB
120 TCAGGCTTCATCCACAGACATCTATATACCAATTTGCAGAGAGCGGTTCTAATGT 179
|||||
122 TCAGGCTTCATCCACAGACATCTATATACCAATTTGCAGAGAGCGGTTCTAATGT 181
DB
180 GCAGTAACTTCAATTCGCTCTGACCTGTTAGTGTATTAACAAAGATGTGAG 239
182 GCAGTAACTTCAATTCGCTCTGACCTGTTAGTGTATTAACAAAGATGTGAG 241
DB
240 CAGAAATTCGCGGCAAAATCCACCGGAGAGATGATTCACATGATGAGATCTCTG 299
242 CAGAAATTCGCGGCAAAATCCACCGGAGAGATGATTCACATGATGAGATCTCTG 301
DB
300 ATATATGTAACCTGTCGACATATATACATCTTCACTGCTGCTCTGACCAACCA 359
302 ATATATGTAACCTGTCGACATATATACATCTTCACTGCTGCTCTGACCAACCA 361
DB
360 TCGTTCTTCCAGATGCTTGAATCAAAATTTGGAGCAAGCGCAAAAGAAAGTCAAG 419
362 TCGTTCTTCCAGATGCTTGAATCAAAATTTGGAGCAAGCGCAAAAGAAAGTCAAG 421
DB
420 TCTACTGGAGATGATTTTCACTATAGATTTGTAATAACACATTACAGAAAGTGA 479
422 TCTACTGGAGATGATTTTCACTATAGATTTGTAATAACACATTACAGAAAGTGA 481
DB
480 GTTAGGAGAGATGTTCTTGTCACTTTTGTCTATGCTATGCTTACCTTCAAGGGCTG 539
482 GTTAGGAGAGATGTTCTTGTCACTTTTGTCTATGCTATGCTTACCTTCAAGGGCTG 541
DB
540 CAATTTCATTTGATTCATCATGATTTTAAATTTGAGTCAAGCTTTTGAAGA 599
542 CAATTTCATTTGATTCATCATGATTTTAAATTTGAGTCAAGCTTTTGAAGA 601
DB
600 TCACAGACAGTTAATGTTTCTGTAACCTTAATTCCTCATATTTTGAAGAAGAGT 659
602 TCACAGACAGTTAATGTTTCTGTAACCTTAATTCCTCATATTTTGAAGAAGAGT 661
DB
660 GATGACGATAA 670
662 GATGACGATAA 672
DB

RESULT 2
CNS000WB/c 526 bp DNA linear GSS 28-JUN-1999
LOCUS Arabidopsis thaliana genome survey sequence T7 end of BAC T12M8 of
DEFINITION

TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
AL094193
AL094193.1 GI:5295347

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Arabidopsis thaliana.
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
JOURNAL

Salanoubat, M., Cholme, N., Artiguenave, F., Brothier, P., Wincker, P., Samsou, D., Saurin, N., Weissenbach, J. and Quetier, F.
Unpublished
2 (bases 1 to 526)
Genoscope.
Direct Submission
Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage ;
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
Web : www.genoscope.cns.fr

FEATURES
source Location/Qualifiers

1. 526
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_id="T12M8"
/clone_id="TAMU"
/note="end : T7"

BASE COUNT 148 a 71 c 90 g 217 t

ORIGIN

Query Match 15.5%; Score 266.6; DB 17; Length 526;

Best Local Similarity 92.1%; Pred. No 8.1e-49; Matches 316; Conservative 0; Mismatches 19; Indels 8; Gaps 3;

1364 AGCCACCATCAACAGACGATACATCCCGAGAAATTCGCTAATAGCTCAGACACAC 1423
526 AGCCACCATCAACAGACGATACATCCCGAGAAATTCGCTAATAGCTCAGACACAC 1473
DB
1424 CACCAACCAACCAACCAACGATGATGCTCCAGTACATCAACCAACCAACATTA 1483
DB
472 CACTACACCAACCAACCAACGATGATGATGATGATGATGATGATGATGATGATG 413
DB
1484 CATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1543
412 CATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1543
DB
1544 CAGCAGACCAAGATTAATTAATGAGAAATTCGCGCTTTATGATACCGATTAATG 1603
352 CAGCAGACCAAGATTAATTAATGAGAAATTCGCGCTTTATGATACCGATTAATG 294
DB
1604 ATGTAACTTAATTCCTCTCTCT-AAAAAATGTTTGAAGCAACAAATTTTAAATG 1662
DB
293 ATGTAACTTAATTCCTCTCTCT-AAAAAATGTTTGAAGCAACAAATTTTAAATG 234
DB
1663 TTAGTGAATTCACATGATTCATTTTAAATTTTAAATTTTAAATTTTAAATTT 1705
233 TTAGTGAATTCACATGATTCATTTTAAATTTTAAATTTTAAATTTTAAATTT 191
DB

RESULT 3
AM038171
LOCUS EST279828 tomato mixed elicitor, RTT Lycopersicon esculentum cDNA
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AM038171
EST279828 tomato mixed elicitor, RTT Lycopersicon esculentum cDNA
AM038171
AM038171.1 GI:5896925
EST.
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;

REFERENCE
D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
Roning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nieman, M.,
Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni,
J.

TITLE
Generation of ESTs from tomato leaf tissue

JOURNAL
Unpublished (1999)

COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
1. 680
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande Flor"
/db_xref="taxon:4081"
/clone="CLER1N5"
/clone_lib="tomato mixed elicitor, BTI"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cleft - inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTI, jasmonic acid, ethylene, fenthion, ELX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

BASE COUNT
ORIGIN
192 a 120 c 164 g 204 t

Query Match 13.8%; Score 237.8; DB 10; Length 680;
Best Local Similarity 74.1%; Pred. No. 2.1e-42;
Matches 315; Conservative 0; Mismatches 107; Indels 3; Gaps 1;

987 GGTGAGCATCAGAGGTGAGAGCCACCTCTTTCAGAAAGCCCAATTCATCATCTT 1046
1 GTTGAGCCGCTGATGCCAGAAATGTCACCTCTGCAAAAAAGCAATTCCTTCTTCT 60

1047 CACAGATCCAGCCCAATGGCGCTTGAGCAAGTAATGTCTGACCGGAGATAGCGAGATGAA 1106
61 CATGAGGCCACCCATGGCAGCTGAGCAAGTTTATCGAGCCGAGACAGTGAGATGAA 120

1107 GTCCATGAGATGTTGTCAGATTTGAAGATGCCAGATGTCATGACTTTTGTGATGTC 1166
121 GTTGATGATGATGTTGAGATCTTGAAGATGCTTGATGATTTTGTGATGTC 180

1167 AATAAAGTGAAGAAACAATTCATGATCTTTGGAATCGTTTGAAGAAACAAGGTT 1226
181 ACCAAGATGAAGAAACAAGTATGATCTGTGGAATCTTTGTTGAAGAAAGAGGTC 240

1227 ATAGCAGATGTCATATCTCTTGGCAGTGAAGCATTTTCAAGATTTTACAGAAAGAG 1286
241 TTGGCAGATGTCATATCCCTTGGCAGTGTGAGCCCTTTCAAGTCGATGTCGAGAG 300

1287 TTGCACCGTTACTCA---TCATCTCTTGGTGTGAGATTTGTTTGAATTAAGTATG 1343
301 TTGGCCCAAGACACCACTATGCAAGGTGTTGAGATTTATGATGTAAGCTGTGG 360

1344 AACCATGACTGTGACTCAGACCAACATCAACACTGCAATACCATCCCGCAATATGTC 1403
361 AACCATGGCTGTGATGCGGCTGACAAATTAACATTTGTAACCTAATATTAAGCAGTTC 420

1404 CGTAA 1408
421 CAAAA 425

RESULT 4
BO505017/c 769 bp mRNA linear EST 22-JUL-2002

DEFINITION
EST612432 Generation of a set of potato cDNA clones for microarray
analyses mixed potato tissues Solanum tuberosum cDNA clone STMGB40
3' end, mRNA sequence.

ACCESSION
BO505017

VERSION
BO505017.2 GI:21920982

KEYWORDS
EST.

SOURCE
Solanum tuberosum
potato.

REFERENCE
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, M., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamycheva, S.A.

AUTHORS
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Jun 10, 2002 this sequence version replaced gi:21363886.
Other ESTs: EST612431

JOURNAL
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@igf.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cdna@resgen.com
Seq primer: 17.

FEATURES
source
1. 769
/organism="Solanum tuberosum"
/cultivar="Kennebec or Blunje"
/db_xref="taxon:4113"
/clone="STMGB40"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
tubers, or roots."

BASE COUNT
ORIGIN
239 a 176 c 140 g 214 t

Query Match 13.5%; Score 232.4; DB 14; Length 769;
Best Local Similarity 73.0%; Pred. No. 3.4e-41;
Matches 327; Conservative 0; Mismatches 116; Indels 5; Gaps 2;

961 CTGCTACTACTAAGCAAGAAATATGCTGAGCGATCAGAGGCTGAAAGCCACCTACTTC 1020
747 CAGTTCGCCAAGTCAGAAAGATATATCATGTTAGAGGCTCTATCCCAAAATTCGTGCTCC 688

1021 TTCAGAAAGCCCAATTCATTCATTCACAGAGTCCAGCCAAATGGCGTTGAGCAAGTAA 1080
687 TCCAAAAAAGGCAA-TCTTTCATTCATCAATGAGGCCAGCCATGCGACTGAGCAAGTTT 629

1081 TGTCTGACCGGATAGCAGATGAAGTGAATGATGATGATGATGATGATGATGATGATG 1140
628 TATCAGATGCAACACAGTGAAGTGAAGTGAATGATGATGATGATGATGATGATGATG 569

1141 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
568 GGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 509

1201 ACTGTTTGAAGAAAAAAGAGTTATAGCAGATGTCATATCTTGGGAGTGTGAAG 1260
508 ACTCATTTGTTAGAAAGCAAGAGGTGTCAGATGTCATATCCCTTGGCAGTGTGAGG 449

1261 CATTTCAAGATTTTACGAAAGAGTTGACACCGTTAC---TCATCACTCTTGTGTTT 1316
448 CCTTTCAAGCTGATGATTCAGATGTTTGGCCAAAGCAGACGACCTGTTGAGAGTTGT 389

QY 1317 TGGAGATTGTTTTCATTAACATGACACGAGCTTGTGACATGACCAACCATCAAC 1376
 Db 388 TGGAGATTGTTTTCATTAACATGACACGAGCTTGTGACATGACCAACCATCAAC 329
 QY 1377 AACGCAATACCATCTCGAGAAATGGC 1404
 Db 328 AATGTAACTTAATTTAGACAGTCC 301

RESULT 5
 AY104964 1128 bp mRNA linear HNC 25-MAY-2002
 LOCUS Zea mays PCO105783 mRNA sequence.
 DEFINITION AY104964
 ACCESSION AY104964
 VERSION AY104964.1 GI:21208042
 KEYWORDS HNC.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 1128)
 Hailey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsett, M.S.,
 Arthur, J.W., Hailey, M., Morgante, M., and Tingey, S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1128)
 Coe, E.C.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 Location/Qualifiers
 1..1128
 /organism="Zea mays"
 /db_xref="MaizeDB:636128"
 /db_xref="taxon:4577"
 /clone="PCO105783"
 /clone_id="Maize Mapping Project/Dupont Consensus
 library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 TITLE
 JOURNAL
 FEATURES
 source
 BASE COUNT 355 a 199 c 242 g 332 t
 ORIGIN
 Query Match 13.1%; Score 226; DB 11; Length 1128;
 Best Local Similarity 57.3%; Pred. No. 9.4e-40;
 Matches 527; Conservative 0; Mismatches 330; Indels 63; Gaps 4;

QY 485 GGAGGATGTTGCTTCATTCATTCCTATGCTATGCTGAGTGTCAAGGGGCTGCAAT 544
 Db 11 GGAAGATTTCCTGCTTCATTCCTATGCTATGCTGAGTGTCAAGGGGCTGAGAG 70
 QY 545 TCAATGATTCATTCATTCCTATTCATTCATTCATTCATTCATTCATTCATTCATTC 604
 Db 71 CCAATTAACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 130
 QY 605 GACAGTTAATGTTCTGTAATACTTAATTCCTCATATTTGAG--GAAGAAGAGTGA 661
 Db 131 GGTGTTAATGTTCTGTAATACTTAATTCCTCATATTTGAG--GAAGAAGAGTGA 190
 QY 662 TGACGATTAATGTTCTGTAATACTTAATTCCTCATATTTGAG--GAAGAAGAGTGA 721
 Db 191 TGATCAAGGATTAATTCATTCCTCATATTCCTCATATTCCTCATATTCCTCATATTC 250
 QY 722 TGGCAGAAATTAACACGAGAGACTTAAGATGCTTTTACGCTTGAGTCCACCAAGTTT 781
 Db 251 GAACACAAATGAGAAATTAACAGC--ATGTACACTCAACATTAATGAGATCAAGTTTACC 307

QY 782 AACTAATGACAGAAATATGATACCCCTACTTAATGATGAAACCGTGTGAGATA 841
 Db 308 TGAAGATGAGGAGAGATCTGAGAGACAACTTGTGCAAGGGGAGAAATGGAGCTTCTTAC 367
 QY 842 TCCCGAGGACAGAGAGCTTGTGAGCAATTTGAGATGACAGCAATTCACACAGCAT 901
 Db 368 AATGCTGATGATTCCTGCTCAATCTTTACATGAGAGAACTTTACACCAACAGAT 427
 QY 902 AGCCACCTCTCTGAGAGCGTGGCTAAAGTTATATATACAAAGCAAGCTGTGTCC 961
 Db 428 ACTAGAGTTGGG----- 440
 QY 962 TGTCTACTAGACAGAAAGTTATCTGCTGAGCGATGAGAGGCTAGAAAGCACTACTTCT 1021
 Db 441 -----AAGACAGGAAGCTAT--CTGAGAGATCTGAGCGCTAGAAATCGGCACTCTCT 490
 QY 1022 TCAGAAAGCCCAATTCATTCATTCGACAGAGTCCAGCAATGCGCTGAGCAAGTAT 1081
 Db 491 GCAAAAGCAGAGCTTCTTCATTCGACAGGCGGACCACTCACTGAGCAAGTAT 550
 QY 1082 GTCTGACCGGATGACGAGAGATGAAGTGCATGACAGATTTGCAAGATCCCA 1141
 Db 551 CTGCGACCGTGAATGATGAAGATGAAGTGAATGATGATGATGATGATGATGATGATG 610
 QY 1142 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1201
 Db 611 AATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
 QY 1202 CTCGTTGTAAGAAACAAAGGTTATGACAGATGCTATCTTGGGACATGTAAGC 1261
 Db 671 TTCAATTTGTAAGAAACAAAGGTTATGATGATGATGATGATGATGATGATGATGATG 730
 QY 1262 ATTTTCAATTTTGAAGAAAGATGACACCGTCTACTCATCTCTCTGCTGTTGGAG 1321
 Db 731 ATTCTCCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 790
 QY 1322 ATGCTTTGATTAACATGTAAGCAATGACAGCTGTCATGACAGCAATCAACAGC 1381
 Db 791 TTTCCTCATGATTAACATTTGTAAGCAATGATGATGATGATGATGATGATGATGATG 850
 QY 1382 CAATACCATCTCGAGAAAT 1401
 Db 851 CAATACCATCTCGAGAAATTT 870

RESULT 6
 BG648271 846 bp mRNA linear EST 24-APR-2001
 LOCUS EST509890 HOGA Medicago truncatula cDNA clone pROGA-19E20 5' end,
 mRNA sequence.
 DEFINITION BG648271
 ACCESSION BG648271.1 GI:13783383
 VERSION
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
 Medicago.
 1 (bases 1 to 846)
 Hahn, M.G., Ojane-Reuhs, T., Samac, D., Town, C.D., Van Aken, S.,
 Uterback, T., Cho, J., and Fraser, C.M.
 ESTs from roots of Medicago truncatula treated with
 oligogalacturonides of DP 6-20
 Unpublished (2001)
 Contact: Michael G. Hahn
 Complex Carbohydrate Research Center
 University of Georgia
 220 Riverbend Road, Athens, GA 30602-4712, USA
 Tel: 706-542-4457
 Fax: 706-542-4412
 Email: hahn@ccrc.uga.edu
 G391877e TIGR sequence name: MTMCJ34TK More information is

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rst

Page 8

VERSION AV817525.1 GI:19859415
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
REFERENCE Rosidae; eukotsid II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamaya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawaji, J., Itoh, M., Ishi, Y., Aikawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinozaki, K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rcc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FliC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified Bluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

SOURCE

Location/Qualifiers
1..391
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFLO9-94-M06"
/clone_lib="RAFLO9"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
BASE COUNT 102 a 100 c 73 g 116 t
ORIGIN

Query Match 11.5%; Score 197.8; DB 10; Length 391;
Best Local Similarity 71.7%; Pred. No. 1.5e-33;
Matches 259; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

OY 1051 GAGTCAGCCCAATGGCGCTTGAGCAATATGCTGACCGGATAGCGAGATGAAGTCG 1110
DB 388 GAGCTCAGCCCAATGGCGCTTGAGCAATATGCTGACCGGATAGCGAGATGAAGTCG 329
OY 1111 ATGACGATGTTGCAATTTTGAAGATCCGAGATGCTTGATGATGTTGATGATGATA 1170
DB 328 ATGATGATGTTGCAATTTTGAAGATCCGAGATGCTTGATGATGTTGATGATGATA 269
OY 1171 AAGATGAAAGCAATTCATGATCTTTGGAACCTGTTGTAAGAAACAAAGGTTTAG 1230
DB 268 AAGATGAAAGCAATTCATGATCTTTGGAACCTGTTGTAAGAAACAAAGGTTTAG 209
OY 1231 CAGATGTCATATCTTGGGATGTAAGCAATTTTCAAGATTTTACGAGAAAGATTTAG 1290
DB 208 CAGATGTCATATCTTGGGATGTAAGCAATTTTCAAGATTTTACGAGAAAGATTTAG 149
OY 1291 ACCGTTACTCATCTCTTGGTGTGAGATGTTTGAATGATTTGATTAAGTATGAAACATG 1350
DB 148 TTGGAACACCGCACTGATTTGCTGAGAGTGTGATTAAGTATGAAACATGTTGAAACACG 89
OY 1351 GACTGTGATGACCGCACTGATTTGCTGAGAGTGTGATTAAGTATGAAACATGTTGAAACACG 1410
DB 88 GCTGTGATGACCGCACTGATTTGCTGAGAGTGTGATTAAGTATGAAACATGTTGAAACACG 29
OY 1411 G 1411

Db 28 G 28
RESULT 12
BO740672
LOCUS
DEFINITION
Sac50505.y1 Gm-cl076 glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl076-3274 5', similar to FR:09ZNT9 09ZNT9
FERTILIZATION INDEPENDENT SEED 2 PROTEIN. ;, mRNA sequence.
ACCESSION BO740672
VERSION BO740672.1 GI:21887455
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
REFERENCE Glycine max
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800) 533-4363 or contact: custresgen.com web site: www.resgen.com
Seq primer: 409P from Glibco
High quality sequence stop: 426.
Location/Qualifiers
1..561
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl076-3274"
/clone_lib="Gm-cl076"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="DH10B"
/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Strategene Bluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Yalla Vodkin lab University of Illinois)."

FEATURES

SOURCE

BASE COUNT 166 a 91 c 123 g 181 t
ORIGIN

Query Match 11.0%; Score 190; DB 14; Length 561;
Best Local Similarity 72.2%; Pred. No. 8.6e-32;
Matches 247; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

OY 1051 GAGTCAGCCCAATGGCGCTTGAGCAATATGCTGACCGGATAGCGAGATGAAGTCG 1110
DB 388 GAGCTCAGCCCAATGGCGCTTGAGCAATATGCTGACCGGATAGCGAGATGAAGTCG 329
OY 1111 ATGACGATGTTGCAATTTTGAAGATCCGAGATGCTTGATGATGTTGATGATGATA 1170
DB 328 ATGATGATGTTGCAATTTTGAAGATCCGAGATGCTTGATGATGTTGATGATGATA 269
OY 1171 AAGATGAAAGCAATTCATGATCTTTGGAACCTGTTGTAAGAAACAAAGGTTTAG 1230
DB 268 AAGATGAAAGCAATTCATGATCTTTGGAACCTGTTGTAAGAAACAAAGGTTTAG 209
OY 1231 CAGATGTCATATCTTGGGATGTAAGCAATTTTCAAGATTTTACGAGAAAGATTTAG 1290
DB 208 CAGATGTCATATCTTGGGATGTAAGCAATTTTCAAGATTTTACGAGAAAGATTTAG 149
OY 1291 ACCGTTACTCATCTCTTGGTGTGAGATGTTTGAATGATTTGATTAAGTATGAAACATG 1350
DB 148 TTGGAACACCGCACTGATTTGCTGAGAGTGTGATTAAGTATGAAACATGTTGAAACACG 89
OY 1351 GACTGTGATGACCGCACTGATTTGCTGAGAGTGTGATTAAGTATGAAACATGTTGAAACACG 1410
DB 88 GCTGTGATGACCGCACTGATTTGCTGAGAGTGTGATTAAGTATGAAACATGTTGAAACACG 29
OY 1411 G 1411

BASE COUNT	ORIGIN
136 a	87 c 95 g 149 t

Query Match	11.0%;	Score 188.8;	DB 12;	Length 467;
Best Local Similarity	67.0%;	Pred. No. 1.5e-31;		
Matches 300;	Conservative 0;	Mismatches 142;	Indels 6;	Gaps 2;

Dd
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 TACAGCTGATGAAGAAGTCTTTTGATTATTAATTCGAAACCCTGTGAATTCCTACCAATATTCCTC 60

Db 61 ATGTCGTCTCTTCACAAATCCCTACTTTTCTTAAGAGAGTGTGGCGCTATAAATAAAG 120

Db 121 CAAAGCAAAAAAGAGGAGTTGCCG--GCAGGAATTGTGGTTTTCACACTATAGGAGACTGTT 177

Db 178 ACAATGCCCTCGAAGACTGAAGTGAAGACTTTTCTGTGCATTTTGGCTGATGC 237
 Dy 517 TATGTGGTAGCTTCAAGGGGCTCAATTTCAATTTGAATTCATCATCATGATTTTATTTGAAT 576

DB 238 AGTGTGCAGGCTTTAAGGGTTTGGCATTTTCATCTTTGTGCATCATCATGATCTATTCAACT 297

DY 577 TTGATGTCAGGCTTTTCAGAGATATCCAGACAGCTTAAGTTTTCGTAATAACTTATTCCT 636

DD 258 TTGAGTCTGGGTCACCAATGATATACCAAGCAGTGAACGCTCTGTGAAAAACATGACATAI 357
 QY 637 TCATATTGAGGAAGAAGGAAGTGTATGACAT---AAATTTGAGCCCTTCTCTCTCTGCT 693

694 CGAACTCTGTAAGCGGAGACAAAGAGG 721

RESULT 14
AT164598

ACCESSION	A1164598	504 bp	linear	531	03-DEC-1997
DEFINITION	Hybrid aspen plasmid library tremula x tremulaoides CDNA 5', mRNA sequence.				
LOCUS	A1164598	504 bp	linear	531	03-DEC-1997
KEYWORDS	Hybrid aspen plasmid library tremula x tremulaoides CDNA 5', mRNA sequence.				

KEYWORDS	EST.
SOURCE ORGANISM	Populus tremula x Populus tremuloides.
	Populus tremula x Populus tremuloides

REFERENCE
 Althoff, R. 1969. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 1 (bases 1 to 364).
 Althoff, R. 1970. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 2 (bases 365 to 640).
 Althoff, R. 1971. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 3 (bases 641 to 916).
 Althoff, R. 1972. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 4 (bases 917 to 1192).
 Althoff, R. 1973. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 5 (bases 1193 to 1468).
 Althoff, R. 1974. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 6 (bases 1469 to 1744).
 Althoff, R. 1975. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 7 (bases 1745 to 2020).
 Althoff, R. 1976. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 8 (bases 2021 to 2296).
 Althoff, R. 1977. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 9 (bases 2297 to 2572).
 Althoff, R. 1978. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 10 (bases 2573 to 2848).
 Althoff, R. 1979. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 11 (bases 2849 to 3124).
 Althoff, R. 1980. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 12 (bases 3125 to 3400).
 Althoff, R. 1981. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 13 (bases 3401 to 3676).
 Althoff, R. 1982. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 14 (bases 3677 to 3952).
 Althoff, R. 1983. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 15 (bases 3953 to 4228).
 Althoff, R. 1984. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 16 (bases 4229 to 4504).
 Althoff, R. 1985. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 17 (bases 4505 to 4780).
 Althoff, R. 1986. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 18 (bases 4781 to 5056).
 Althoff, R. 1987. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 19 (bases 5057 to 5332).
 Althoff, R. 1988. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 20 (bases 5333 to 5608).
 Althoff, R. 1989. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 21 (bases 5609 to 5884).
 Althoff, R. 1990. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 22 (bases 5885 to 6160).
 Althoff, R. 1991. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 23 (bases 6161 to 6436).
 Althoff, R. 1992. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 24 (bases 6437 to 6712).
 Althoff, R. 1993. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 25 (bases 6713 to 6988).
 Althoff, R. 1994. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 26 (bases 6989 to 7264).
 Althoff, R. 1995. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 27 (bases 7265 to 7540).
 Althoff, R. 1996. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 28 (bases 7541 to 7816).
 Althoff, R. 1997. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 29 (bases 7817 to 8092).
 Althoff, R. 1998. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 30 (bases 8093 to 8368).
 Althoff, R. 1999. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 31 (bases 8369 to 8644).
 Althoff, R. 2000. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 32 (bases 8645 to 8920).
 Althoff, R. 2001. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 33 (bases 8921 to 9196).
 Althoff, R. 2002. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 34 (bases 9197 to 9472).
 Althoff, R. 2003. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 35 (bases 9473 to 9748).
 Althoff, R. 2004. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 36 (bases 9749 to 10024).
 Althoff, R. 2005. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 37 (bases 10025 to 10300).
 Althoff, R. 2006. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 38 (bases 10301 to 10576).
 Althoff, R. 2007. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 39 (bases 10577 to 10852).
 Althoff, R. 2008. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 40 (bases 10853 to 11128).
 Althoff, R. 2009. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 41 (bases 11129 to 11404).
 Althoff, R. 2010. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 42 (bases 11405 to 11680).
 Althoff, R. 2011. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 43 (bases 11681 to 11956).
 Althoff, R. 2012. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 44 (bases 11957 to 12232).
 Althoff, R. 2013. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 45 (bases 12233 to 12508).
 Althoff, R. 2014. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 46 (bases 12509 to 12784).
 Althoff, R. 2015. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 47 (bases 12785 to 13060).
 Althoff, R. 2016. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 48 (bases 13061 to 13336).
 Althoff, R. 2017. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 49 (bases 13337 to 13612).
 Althoff, R. 2018. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 50 (bases 13613 to 13888).
 Althoff, R. 2019. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 51 (bases 13889 to 14164).
 Althoff, R. 2020. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 52 (bases 14165 to 14440).
 Althoff, R. 2021. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 53 (bases 14441 to 14716).
 Althoff, R. 2022. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 54 (bases 14717 to 14992).
 Althoff, R. 2023. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 55 (bases 14993 to 15268).
 Althoff, R. 2024. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 56 (bases 15269 to 15544).
 Althoff, R. 2025. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 57 (bases 15545 to 15820).
 Althoff, R. 2026. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 58 (bases 15821 to 16096).
 Althoff, R. 2027. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 59 (bases 16097 to 16372).
 Althoff, R. 2028. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 60 (bases 16373 to 16648).
 Althoff, R. 2029. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 61 (bases 16649 to 16924).
 Althoff, R. 2030. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 62 (bases 16925 to 17200).
 Althoff, R. 2031. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 63 (bases 17201 to 17476).
 Althoff, R. 2032. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 64 (bases 17477 to 17752).
 Althoff, R. 2033. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 65 (bases 17753 to 18028).
 Althoff, R. 2034. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 66 (bases 18029 to 18304).
 Althoff, R. 2035. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 67 (bases 18305 to 18580).
 Althoff, R. 2036. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 68 (bases 18581 to 18856).
 Althoff, R. 2037. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 69 (bases 18857 to 19132).
 Althoff, R. 2038. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 70 (bases 19133 to 19408).
 Althoff, R. 2039. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 71 (bases 19409 to 19684).
 Althoff, R. 2040. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 72 (bases 19685 to 19960).
 Althoff, R. 2041. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 73 (bases 19961 to 20236).
 Althoff, R. 2042. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 74 (bases 20237 to 20512).
 Althoff, R. 2043. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 75 (bases 20513 to 20788).
 Althoff, R. 2044. *Die Kieferpflanzen der Erde*. Gustav Fischer Verlag, Stuttgart. 76 (bases 20789 to 21064).

TITLE	Author	Year
Gene discovery in the wood-forming tissues of poplar: Analysis of	M. Smith, B. Lundberg, C. Larsson, T. Boerjan, W. Gustafsson, P. M. Sandberg, G. Olsson, O. Teerilä, T. T. Boerjan, W. Gustafsson, P. Uhlen, M. Sundberg, B. and Lundberg, J.	2007

JOURNAL
PROC. Natl. Acad. Sci. U.S.A. 95 (22), 13330-13335 (1998)
MEDLINE
99007314
COMMENT
Contact: Sterky F
Department of Biotechnology

Royal Institute Of Technology (KTH)
Teknikringen 34, S-100 44 STOCKHOLM, Sweden
Tel: +46 8 790 8287
Fax: +46 8 24 54 52

PCR Primers

FORWARD: AAGGGGATGTCGCAAGCGG
 BACKWARD: GCTCCGGCTGTAATGTTGTCG
 Seq primer: GGTGTAAAGCAGCGCCG
 High quality sequence stop: 364.

FEATURES

source

1..364

/organism="Populus tremula x Populus tremuloides"

/db_xref="taxon:47664"

/clone_lib="Hybrid aspen plasmid library"

/tissue_type="Cambial region"

/dev_stage="1.5 m actively growing tree"

/lab_host="E. coli"

/note="Vector: pBluescript SK; Site_1: SalI; Site_2: NotI; Cambial region tissues, including developing xylem, the meristematic cambial zone and the developing and mature phloem, was harvested from 1.5 m actively growing trees. cDNA was prepared and cloned into lambda gt22a. DNA was isolated and subcloned into pBluescript SK using SalI and NotI restriction enzymes."

BASE COUNT

111 a 68 c 94 g 90 t 1 others

ORIGIN

Query Match 10.2%; Score 176.4; DB 9; Length 364;
 Best Local Similarity 75.5%; Pred. No. 8.5e-29;
 Matches 244; Conservative 0; Mismatches 77; Indels 2; Gaps 2;

951 GCTGTGTCCTCCCTACTAAGACAGAAAGTTATCTGTCAGCCGATCAGAGCTAGAGC 1010
 44 GCCATCTACATATTTCCAAAGACAGAAATTTATCAATTTGAACGCGCTGATGAGAAAC 103
 1011 CACCTACTCTTCAGAAAGCCCAATTTATCTATCTCAGACAGATCCCAATGCGCCTT 1070
 104 CGTACACTCTCTTCACAAAGACAAATTTTTCACCTACATAGAGCCAGCCAGTCACT 163
 1071 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGAGTGAAGTCAGATGTTGATTTT 1130
 164 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGAGTGAAGTGAAGTGAAGTGAAGT 223
 1131 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGAGTGAAGTGAAGTGAAGTGAAGT 1190
 224 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGAGTGAAGTGAAGTGAAGTGAAGT 282
 1191 CATCTTGAACCGCTTGTATAGAAAGGCTTATAGAGATGCTGATTTCTTGG 1250
 283 CACTTGTGACATCATTTTGTGAGAGAGCGGCTTGTGAGATGACATATC-CCGGG 341
 QY 1251 GCATGTGAAGCATTTTCAAGATT 1273
 Db 342 GCATGTGAAGCATTTTCAAGATT 364

RESULT 15
 BG314044 405 bp mRNA linear EST 16-Apr-2001
 LOCUS BG314044
 DEFINITION WHE2467.D04_G0725 Trifolium monoccoccum early reproductive apex cDNA
 library Trifolium monoccoccum cDNA clone WHE2467_D04_G07, mRNA
 sequence.
 accession BG314044.1 GI:13115847
 version BG314044
 keywords EST.
 source Trifolium monoccoccum
 organism Trifolium monoccoccum

REFERENCE
 AUTHORS Anderson O.D., Chao S., Dubcovsky J., Echenique V., Han P.S., Hsia
 1 (bases 1 to 405)
 TITLE Trifolium monoccoccum
 JOURNAL Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; Pooidae
 1 (bases 1 to 405)
 C.C., Kang Y., Lazo G.R., Miller R., Kausch C.J., Seaton C.L.,
 Stamova B. and Tong J.C.
 The structure and function of the expressed portion of the wheat
 genomes - Early reproductive apex cDNA library from Trifolium
 monoccoccum
 unpublished (2001)

COMMENT

Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific
 West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: oanderson@wrs.ars.gov

FEATURES

source

1..405

/organism="Trifolium monoccoccum"

/db_xref="taxon:4568"

/clone_lib="WHE2467_D04_G07"

/tissue_type="Early reproductive apex"

/dev_stage="Seven week-old plants"

/lab_host="E. coli XL0R"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and
 poly(A) RNA were prepared from apex at double-ridge stage
 to terminal-spikelet stage during transition from
 vegetative state to flower state, a cDNA library was made,
 and the cDNA clones were in vivo excised at the
 University of California, Davis (V. Echenique, B. Stamova,
 J. Dubcovsky). Plasmid DNA preparations and DNA
 sequencing were performed in the OD Anderson lab (all
 other authors)."

BASE COUNT

105 a 85 c 104 g 111 t

ORIGIN

Query Match 10.1%; Score 174.6; DB 12; Length 405;
 Best Local Similarity 68.8%; Pred. No. 2.2e-28;
 Matches 240; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

1056 CAGCCAAATGGCGCTTGAAGCAATGTTCTGACCGGATAGGAGATGAATCATGATGAC 1115
 4 CAGCCAAATGGCGCTTGAAGCAATGTTCTGACCGGATAGGAGATGAATCATGATGAT 63
 1116 GATGTTGCAATTTTGAAGATGCCAGATCTTGATGATTTGTGATGATGAATAAGT 1175
 64 GACATTTGCGACTTGAAGATGACGAGATCTTGATGATTTGTGATGATGATGATGAT 123
 QY 1176 GAAAGCAATTCATCTTTGAACTGTTGTAGAAACAAAGGCTTATGACAGT 1235
 Db 124 GAGAAAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
 QY 1236 GGTCAATATCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295
 Db 184 GGGCAATATATCTTGGGCTTGAAGGCTTCTCCCGGCTCATGAGACCGCATGATGAT 243
 QY 1296 TACTATATCTTCTGTTGTTGAGATTTTGTGATTAATATATGATGATGATGATGAT 1355
 Db 244 AACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
 QY 1356 GTGCACTAGCCAGCATCAACCAATGATCAATCAATGATGATGATGATGATGATGAT 1404
 Db 304 CTGATGATGCGCGCGCATGACATGACATGACATGACATGACATGACATGACATGAC 352

Search completed: June 19, 2003, 22:08:41
 Job time : 2572 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 20:10:48 ; Search time 1286 Seconds

(without alignments)
9309.105 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722
Sequence: 1 caagctctctcaatttgc.....aatgattctgtataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7816032 seqs, 3476047932 residues

Total number of hits satisfying chosen parameters: 15632064

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PC1_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*
13: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366.8	21.3	1722	US-10-425-114-11348	Sequence 11348, A
2	366.8	21.3	1838	US-10-424-599-111422	Sequence 111422, A
3	257.4	14.9	1841	US-10-425-114-31883	Sequence 31883, A
4	236.2	13.7	2640	US-10-231-778-6	Sequence 6, Appl1
5	230.2	13.4	1004	US-10-437-963-11609	Sequence 11609, A
6	227.2	13.2	1070	US-10-424-599-14113	Sequence 14113, A
7	192.4	11.2	716	US-10-424-599-111421	Sequence 111421, A
8	185.6	10.8	443	US-09-837-604B-13507	Sequence 13507, A
9	185.6	10.8	443	US-09-837-604A-13507	Sequence 13507, A
10	174	10.1	558	US-10-389-048-21629	Sequence 21629, A
11	158.8	9.2	2700	US-10-437-963-91042	Sequence 91042, A
12	138.8	8.1	554	US-09-531-113-5697	Sequence 5697, A
13	138.8	8.1	554	US-09-531-113-5697	Sequence 5697, A
14	125.6	7.3	322	US-10-424-599-40645	Sequence 40645, A
15	104	6.0	453	US-09-837-604B-6414	Sequence 6414, A
16	104	6.0	453	US-09-837-604A-6414	Sequence 6414, A
17	104	6.0	453	US-10-437-963-97425	Sequence 97425, A
18	101.2	5.9	2439	US-10-425-114-30709	Sequence 30709, A
19	100	5.8	531	US-09-837-604B-940	Sequence 940, A
20	100	5.8	531	US-09-837-604A-940	Sequence 940, A

c	21	91.8	5.3	632	10	US-10-424-599-73456	Sequence 73456, A
c	22	82.8	4.8	250	5	US-09-531-113-40606	Sequence 40606, A
c	23	82.8	4.8	250	6	US-09-531-113-40606	Sequence 40606, A
c	24	71	4.1	274	10	US-10-389-048-16407	Sequence 16407, A
c	25	69	4.0	921	10	US-10-424-599-108137	Sequence 108137, A
c	26	63.8	3.7	201	13	US-60-466-412-340472	Sequence 340472, A
c	27	63.8	3.7	56447	13	US-60-466-412-86674	Sequence 86674, A
c	28	58.2	3.4	939	10	US-10-437-963-81637	Sequence 81637, A
c	29	57.8	3.4	6458	9	US-10-231-778-7	Sequence 7, Appl1
c	30	57.2	3.3	241	10	US-10-424-599-52847	Sequence 52847, A
c	31	57.2	3.3	492	9	US-10-144-771-38577	Sequence 38577, A
c	32	56.8	3.3	961	9	US-10-144-771-31796	Sequence 31796, A
c	33	54.2	3.1	8524	9	US-10-311-455-1814	Sequence 1814, A
c	34	53.2	3.1	2892	1	PCR-US01-29288A-255	Sequence 255, App
c	35	53.2	3.1	2892	10	US-10-149-310-255	Sequence 255, App
c	36	53	3.1	350	10	US-10-424-599-3668	Sequence 968, App
c	37	52.2	3.0	8662	9	US-10-240-485-190	Sequence 190, App
c	38	51.6	3.0	867	9	US-10-321-856-340	Sequence 340, App
c	39	51.6	3.0	867	9	US-10-321-856-342	Sequence 342, App
c	40	51.6	3.0	1397	9	US-10-321-856-343	Sequence 343, App
c	41	51.6	3.0	1397	9	US-10-321-856-345	Sequence 345, App
c	42	51.4	3.0	958	9	US-10-144-771-24706	Sequence 24706, A
c	43	51.4	3.0	1178	10	US-10-425-114-16009	Sequence 16009, A
c	44	50.8	3.0	1141	6	US-09-806-708B-22	Sequence 22, Appl1
c	45	50.6	2.9	3673778	9	US-10-312-841-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-425-114-11348

Sequence 11348, Application US/10425114

GENERAL INFORMATION:
APPLICANT: Liu, Jiongong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 11348

LENGTH: 1722

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: 701055296_FLI

Query Match
Best Local Similarity 59.8%; Pred. No. 1.1e+80;

Matches 757; Conservative 0; Mismatches 422; Indels 87; Gaps 5;

QY	228	ACAAATGTGAGCAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	287
DB	161	ACCAATGTGAGCAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	220
QY	288	GAGAAATGTCGAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	347
DB	221	GAGAAATGTCGAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	280
QY	348	CTAGGCAACCAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	407
DB	281	CTTCAAAATGTCGAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	340
QY	408	AGAAATGTCGAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	467
DB	341	AGAGAGTGTGAGAAATGTCGCGGAATCTCAACGAGAAATGATTTCAACTGAT	397


```

Db      912 AATGAGGAGGAAAAATTTGGTCTGATCATCTGGACCAATGACAACTGGAGATGTG 971
OY      873 GAGATGACCAAGCAACATTCACCA-----GCCATAGCCCACTCTTCTGAGCGCT 923
Db      972 GAATCCAGTTTCAACATTCACAGGTGTTTCAATTCGATGCCCAATCTTCTGAGACCT 1031
OY      924 GGTGCTAAAGTTATATTTGACAGCAAGGTGTGTCCTGCTACTAAGACAGAAAGTTA 983
Db      1032 GAATGTAGTAATCAATATGTAATAAGTATCCGCTGCTGCTTAAACAAAGAGCTA 1091
OY      984 TCTGCTGAGCGATCAGAGGCTGAGACCCACTACTCTTCTGAGAAAGCCCAATCTCAT 1043
Db      1092 AGCATGTGATCTCAGCACTCAGAAACCCGAAATGCTCTGAGAAAGAGACTTTCTTCCAC 1151
OY      1044 TCTCAGAGATCCAGCAATGGCGCTTGAAGCAAGTAATGCTGACCGGGATAGCAGGAT 1103
Db      1152 TCACACAGAGTCCAGCTTATGGCAGCTAGAACAGTGTATCAGACCGTGTACTGAAAC 1211
OY      1104 GAATGATGACGATGTTGACAGATTTTGAAGATCCGAGATGCTGATGACTTTGTGGAT 1163
Db      1212 GAAGTGTAGAGCAATTCAGATCTTGAAGATAGAAGATGCTGACGATTTGTGGAT 1271
OY      1164 GTGAATAAAGATGAAGAAAGCAATTCATGCACTTTGCACTCGTTTGAAGAAACAAAG 1223
Db      1272 GTTTCAGAAAGATGAAGAAAGCAATTCATGCACTTTTATGAGAAAGCAAG 1331
OY      1224 GTTATGACAGATGATCATATCTCTGCGATGTGAGCAATTTCAAGATTTTACGAA 1283
Db      1332 GTGCTGACAGATGATGATATTTCCGAGGCTGTGAGCAATTTTCCAGCTTCAATGAA 1391
OY      1284 GAGTGGACACGTTACTCATCTCTCTGTTGGAGATGTTTGTATTAACATGAG 1343
Db      1392 GAGCGATGATCATCTCCACCTTATTTGGTGTGAGGTTATCATGATCAAACTTGG 1451
OY      1344 AACCATGACTGTGCACTCAGCCACCAATCAACACTGCAATACCTCTCGAAGTTCG 1403
Db      1452 AATCATGATCTTCTGATGCTGTACAAATGAAACACTGATATGATGATGATGAC 1511
OY      1404 CGTAAT 1409
Db      1512 AGAAAT 1517

RESULT 3
US-10-425-114-31883
; Sequence 31883, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 31883
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73226612_FLI
US-10-425-114-31883

Query Match      14.9%; Score 257.4; DB 10; Length 1841;
Best local Similarity 56.5%; Pred. No. 1.8e-53;
Matches 573; Conservative 0; Mismatches 391; Indels 51; Gaps 3;
OY      424 CTGGATGGTACTTTTCAACTATAGAGATGTATATACACATTACGAAAACTGAAGTTA 483
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      493 CTGAAATGACTTTTTTATATACAGTACTACAAATACATACAAAGCTGAAGTTA 552
OY      484 GGGAGATGTGTTTCCATTTGCTCTATGCTATGATGTGATGCTTCAAGGGCTGCAT 543
Db      553 CTGAAATTTTCTGCCCCATTTTGGCTGTGATACATGAGAGCTTCAAGGGCTGAGAT 612
OY      544 TTGATTTGAATTCATCATGATTTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTG 603
Db      613 GGCATTAATATCATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
OY      604 AGACAGTTAATGTTTCTGTAATAACTTAATCTTCAATTTTGAAGAAAGAGATGATG 663
Db      673 AGGCTGTGATGTTATGCTGAAGAGCTGATGCTGGAATAAGCTGATTTGTGGAGAGGAG 722
OY      664 AGCATTAATTTGAGCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
Db      733 TTGATCCAAAGCATCAAACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788
OY      724 GCAGAAATACACAGAGACTTAAAGTATGCTTTTATCCGTTGATTCACCACTTTTAA 783
Db      789 TTGGAAGACACAGCTGAGAAATTCAGGCAATGTTACCCACATTTATGGAATCAGAT 846
OY      784 CTATGAGCACGAAATTCGATACCCCTACTTAATGATGAGAAACCGTGTATGATATC 843
Db      847 -----CACATGAAGATGCCAGGAGGCTGAAAGATG 879
OY      844 CCGAGGCAACAGAGCTTGTGACAAATTTGATGATGACCAACATTTCCACGACCATAG 903
Db      880 ACTATGTGCAAGGAGAAATGGGCTTTCTG-----TAGCAATATCTT 921
OY      904 CCCACTCTTCTCTGAGAGCTGCTGCTAAAGTTATTTGACAAAGCAAGCTGTGCTGCTG 963
Db      922 CAGTTGACCTATGACGCTGATGATGAGAGCAATTTTACCAACAGAGTCTGATAGT 981
OY      964 CTACTAAGACAGAAATTTATCTGTGAGGATCAGAGGCTAGAGCCACCACTTCTTCT 1023
Db      982 TTGGAAGATCAAGAAATCTATGCGAGGATCTGACCCCAAGACCACTGCTTGGC 1041
OY      1024 AGAAAGCCAAATCTATCATCTCCAGAGCTCCAGCAATGGGGCTTGACAGTATATG 1083
Db      1042 AAAAGGTCAGTCTTCATTCACACAGGCGACAGCCAAATGCTTGGAAACAGTCTCT 1101
OY      1084 CTGACCGGATAGCGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143
Db      1102 CAGATCATGATATACAGAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
OY      1144 TGCTTATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
Db      1162 TGCTTCAGCATTTTGTGATGTTACAAAGATGAGAGCTTATATGATGATGATGATGAT 1221
OY      1204 CGTTTGTAAAGAAAGAGGTTATGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1263
Db      1222 CATTTGTGAAAGCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
OY      1264 TTTCAGATTTTACGAAAGAGTTCACCGTACTCATCTCTTGTGTTGGAGAT 1323
Db      1282 TCTCCAGATGCGCATGAGCAACACTGTACAAAACCTGCGCTGTGGGCTGTGGCT 1341
OY      1324 TGTTTTGTATTAATGATGAGACCATGACTTGTGATGATGATGATGATGATGATGATGAT 1383
Db      1342 TCTTCATGATTAACCTGTGAAACCAAGCTTCTTGTGACGCGCCGCGCATGAAATGCTTGA 1401
OY      1384 ATACCATCTCGCAATTTGGCGTAATAGCTCAGACACCAACCAACCAACCAACCAACCA 1438
Db      1402 ACACAAATCTTCAAGGCTACCAAGAGGAGTGGAGCCCAAGTAAATCTACAA 1456

RESULT 4
US-10-231-778-6
; Sequence 6, Application US/10231778
; GENERAL INFORMATION:
; APPLICANT: Billoreau, Pierre
; APPLICANT: Chaudhury, Abdul M.

```



```

APPLICANT: Dennis, Elizabeth S.
APPLICANT: Koltunow, Anna M.G.
APPLICANT: Luo, Ming
APPLICANT: Pascock, William J.
TITLE OF INVENTION: Method for inducing seed development by down-regulating
FILE REFERENCE: 72-98A
TITLE OF INVENTION: expression of the F1S2 gene
CURRENT APPLICATION NUMBER: US/10/231,778
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: 09/398,237
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 60/101,184
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: AU PP6061
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6062
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6063
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6063
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: AU P01345
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: AU P01346
NUMBER OF SEQ ID NOS: 239
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO: 6
LENGTH: 2640
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)-(2439)
OTHER INFORMATION: Nucleotides from 1 to 2439 represent protein
US-10-231-778-6

Query Match
Best Local Similarity 13.7%; Score 236.2; DB 9; Length 2640;
Matches 315; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

954 GTGTCCTCTGCTAAGCAAGAAAGTTATCTGCTGAGCGATCAGAGCTAGAGCCAC 1013
1915 GTGACTGTGTGTGAGAAAGAAAGAGTTACATGACAGCGGTGCGAGCTTAAAGATG 1974
1014 CTACTCTTGAAGAACGCCAATTCATCTCTCAGAGTCCAGCCCAATGGCCCTTGAG 1073
1975 GAGCGTCTTAAAGGTGACAGAGTTCTATCTCCTCCAAACATGACCCCAATGACTTTGAA 2034
1074 CAAGTAATGTCTGACCGGGATGACGAGATGAAAGTGCATGATGTTGCGAGATTGAA 1133
2035 CAAGTAATGTCTAAGCAGATAGCGAGATGAGACTGATGATTTATGCTTTAGATATTAGC 2094
1134 GATGCCAGATGCTTGATGATCTTTGTGATGTAATAAGATGAAAAGCAATTCATCAT 1193
2095 GAAAGCGCTGAGACTTAAAGCTCTTGTGGGTGAGCAAGAGGAAAGCAAGCGTATCAT 2154
1194 CTCTGGAATCTGTTTGAAGAAACAAAGGTTATAGCAGATGGTCTATATCTCTTGGCA 1253
2155 CTCTGGAATCTTATTTTACGAAACAAAGGTTGATGGGATGGACATGTTCTTGGGCA 2214
1254 TGTGAAGCATTTCAAGATTTTACGAAAGAGTTGACCCGTTACTCATCTCTTCTGG 1313
2215 TGTGAAGAGTTTCAAAACCTTCAATAGGAGAGATGAAAGAAATCTTCTCTTCTGATTG 2274
1314 TGTGGAGATGTTTGTGATTAAGTATGAAACCAATGACATGCTTGAGCTAGCCCATC 1373
2275 TGTGGAGATGTTTGTGATTAAGTATGAAACCAATGACATGCTTGAGCTAGCCCATC 2334
1374 AACACTGCAATTCATCTCTC 1394
2335 CACAAATGACATGACATCTC 2355

```

```

RESULT 5
US-10-437-963-11609/C
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Barabarov, Andrey A.
APPLICANT: Barabarov, Andrey A.
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO: 11609
LENGTH: 1004
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-437-963-11609

Query Match
Best Local Similarity 13.4%; Score 230.2; DB 10; Length 1004;
Matches 316; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

950 AGCTGTGTCCTCTGCTAAGCAAGAAAGTTATCTGCTGAGCGATCAGAGCTAGAG 1009
782 AACAGTGTAAAGTTAGTTGGAGAACAGAAAGCTGTGTTGAACAGCTAGCCAGAA 723
1010 CCACCTACTCTTGAAGAACGCCAATTCATCTCTCAGAGTCCAGCCCAATGGCCCT 1069
722 TCGGAGCTCTTCAAAAAGCCAGAGTCTTCTCTCAGAGCTCAACCAATGGCAT 663
1070 TGAGCAGTAAATGCTGACCGGATGACGAGATGAAAGTGCATGATGTTGAGATT 1129
662 GGAGCAGTAAATGCTGACCGGATGACGAGATGAAAGTGCATGATGTTGAGATT 603
1130 TGAAGATGCGCGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1189
602 TGAAGATGCGCGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 543
1190 GCATCTTGAAGTCTGTTTGAAGAAACAAAGGTTATAGCAGATGGTCTATATCTCT 1249
-542 GCATCTTGAAGTCTGTTTGAAGAAACAAAGGTTATAGCAGATGGTCTATATCTCT 483
1250 GGCATGTCAGCAATTTTCAAGATTTTACGAAAGAGTTGACCCGTTACTCATCTCT 1309
482 GGCATGTCAGCAATTTTCAAGATTTTACGAAAGAGTTGACCCGTTACTCATCTCT 423
1310 CTGTGTGAGATGTTTGTGATTAAGTATGAAACCAATGACATGCTTGAGCTAGCC 1369
422 ATGGTGTGAGATGTTTGTGATTAAGTATGAAACCAATGACATGCTTGAGCTAGCC 363
1370 CATCAACAACATGCAATTCATCTCTGAGAAATTCCTGTA 1408
362 CATGAATGCTTCAACAACAATCTTGAAGGCTACTGAA 324

RESULT 6
US-10-424-599-141143/C
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Barabarov, Andrey A.
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO: 11609
LENGTH: 1004
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-437-963-11609

```



```

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141143
; LENGTH: 1070
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1
US-10-424-599-141143

```

```

Query Match      13.2%; Score 227.2; DB 10; Length 1070;
Best Local Similarity 68.6%; Pred. No. 4,9e+46;
Matches 313; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

```

```

QY 949 AAGCTGTGCTCCCTGCTACTAAGACAAAGTATCTGTCAGCGCATCAGAGGCTAGAA 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 824 AAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1009 GCCACCTACTCTTCTGAGAAAGCCAACTTATCATCTCCAGAGTCCAGCCAAATGCGCC 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 764 ACAGTACCTCTTCTGAGAAAGCCAACTTATCATCTCCAGAGTCCAGCCAAATGCGCC 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1069 TTGAGCAAGTAAATGCTGACCGGATAGCGAGATGAAAGTCAAGTGTGAGATTT 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 704 TTGAGCAAGTAAATGCTGACCGGATAGCGAGATGAAAGTCAAGTGTGAGATTT 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1129 TTGAGCAAGTAAATGCTGACCGGATAGCGAGATGAAAGTCAAGTGTGAGATTT 1188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 644 TTGAGCAAGTAAATGCTGACCGGATAGCGAGATGAAAGTCAAGTGTGAGATTT 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1189 TGCATCTTGGAACTGCTTGTGAAGAAACAAAGGTTTACCAATGCTCATCTCTT 1248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 584 TGCATCTTGGAACTGCTTGTGAAGAAACAAAGGTTTACCAATGCTCATCTCTT 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1249 GGGCATGTGAAACATTTTCAAGATTTTACGAGAAAGATTTGACCGTTACTCATCTCT 1308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 524 GGGCATGTGAAACATTTTCAAGATTTTACGAGAAAGATTTGACCGTTACTCATCTCT 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1309 TCTGTGTTGGAGATTTTGTGTTTAACTATGGAACCATGAGTGTGAGCTAGCCCA 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 464 CAGGGTGTGGAATATTTTATGATGCAATTTTACATCATCATGCTCTTCTGATGCTCGGA 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1369 CCATCAACAACATGCAATACCATCTGAGATTTGCC 1404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 404 CCATCAACAACATGCAATACCATCTGAGATTTGCC 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-10-424-599-111421

```

```

; Sequence 111421, Application US/10424599

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: La Rosa Thomas J

```

```

; APPLICANT: Kovalic David K

```

```

; APPLICANT: Zhou Yihua

```

```

; APPLICANT: Cao Yongwei

```

```

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

```

; FILE REFERENCE: 38-21(53223)B

```

```

; CURRENT APPLICATION NUMBER: US/10/424,599

```

```

; CURRENT FILING DATE: 2003-04-28

```

```

; NUMBER OF SEQ ID NOS: 285684

```

```

; SEQ ID NO 111421

```

```

; LENGTH: 716

```

```

; TYPE: DNA

```

```

; ORGANISM: Glycine max

```

```

; FEATURE:

```

```

; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1
US-10-424-599-111421

```

```

Query Match      11.2%; Score 192.4; DB 10; Length 716;
Best Local Similarity 68.5%; Pred. No. 2e-37;

```

```

Matches 281; Conservative 0; Mismatches 126; Indels 3; Gaps 1;
QY 228 AGAATGTGAGCAGATTTGTCGGGAATCTCTACCCGAGAGATGATTTCAACTGAT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 AGCATGTGCGGCAAAATCTCCGGTACCAATGCGGAGAAAGAAATTCAGCTGAT 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 288 GAGAACTCTGATATATTTGTAACCTGCTGCACTATATACATCTTACCTGCTCT 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 GAGAGTCTTTATATTTATTTGCAAGCTGTTGAACTGTATACAAATTTCTTACCGCGCT 333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 CTAGCAACCATCTGTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGCGCAAA 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 CTCAAAATTCCTCTTCTTTAGAGATGTTTGTCTTATATAAATTAAGCAAGCGGTAA 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 408 AGAAGTCAAGATCTTCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 AGAGAGTTAGA---GCAGAAATTTGATTTTCAATTAAGGATCTACCAACATCTT 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 468 CAGAAATCTGAAGTATGAGAGATGTTTCTTCCATTTTGTCTATGCTATGATGATGAT 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 451 CGGAAATCTGAAGTATGAGAGATGTTTCTTCCATTTTGTCTATGCTATGATGATGAT 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 528 TTCAAGGCGCTCAATTTTCAATTTGATTCATCATCATCATCATCATCATCATCATCAT 587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 511 TTAAAGGTTTCTGATTTTCAATTTGATTCATCATCATCATCATCATCATCATCATCAT 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 588 CTTTCGAGATATACCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 571 GTTACTGAAGATTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
US-09-837-604B-13507

```

```

; Sequence 13507, Application US/09837604B

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Byrum, Joseph R.

```

```

; APPLICANT: La Rosa, Thomas J.

```

```

; APPLICANT: Shukla, Hridayabhiman

```

```

; APPLICANT: De la Pena, Robert C.

```

```

; APPLICANT: Bought, Olegs

```

```

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated with

```

```

; TITLE OF INVENTION: Plants

```

```

; FILE REFERENCE: 38-21(51892)B

```

```

; CURRENT APPLICATION NUMBER: US/09/837,604B

```

```

; CURRENT FILING DATE: 2001-04-18

```

```

; PRIOR APPLICATION NUMBER: US 60/197,872

```

```

; PRIOR FILING DATE: 2000-04-19

```

```

; NUMBER OF SEQ ID NOS: 81288

```

```

; SEQ ID NO 13507

```

```

; LENGTH: 443

```

```

; TYPE: DNA

```

```

; ORGANISM: Oryza sativa nipponbare

```

```

; FEATURE:

```

```

; OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
US-09-837-604B-13507

```

```

Query Match      10.8%; Score 185.6; DB 5; Length 443;
Best Local Similarity 72.0%; Pred. No. 8.3e-36;
Matches 242; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```

```

QY 950 AGCTGTGCTCCCTGCTACTAAGACAAAGATTTCTGTCGAGCTAGAGCTAGAG 1009
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 AACAGTGTACAGTTTGGGAAACAAAGAGCTGTCTTGAACAGAGCTGATCCAGAAA 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1010 CCACCTACTTCTTCAAGAAAGCCAAATTTATATCTCTACAGATGTCACCAATGGCGCT 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 158 TCGGAGCTCTTCAAAACAGCCAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1070 TGAGCAAGTATGCTGACCGGAGATGAGGAGATGAGGAGTCAAGTCAAGTGTGAGATTT 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 218 GAGCAAGTATGCTGACCGGAGATGAGGAGATGAGGAGTCAAGTCAAGTGTGAGATTT 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1130 TGAAGTCCGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

; TITLE OF INVENTION: Them, and Methods for Using Them
; FILE REFERENCE: 11000.1041u1c1
; CURRENT APPLICATION NUMBER: US/10/389,048
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 25129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21629
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
;
Db 278 TGAAGATGAGAAAGATGCTGATGATTTTGTGATGTTACAAAAGCAGAACTTATAT 337
Oy 1190 GCATCTTGGAACTCGTTTGTAGAAAAGAGGTTATAGCAGATGTCATATCTTGT 1249
Db 338 GCATATGTGAAATTCATTTCTGCGAAACAAAGGGTACTAGCGGATGCGCATATTCCTG 397
Oy 1250 GGCATGTGAAGCATTTTTCAGATTTTACGAGAAAGA 1285
Db 398 GGCATGCGAAGCATTTTCTCGATTTTCATGACAGAGA 433

```

RESULT 9.

```

US-09-837-604A-13507
; Sequence 13507 Application US/09837604A
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayadhiranjan
; APPLICANT: De La Pena, Robert C.
; APPLICANT: Bougri, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3432-029-p1-K1-H2
US-09-837-604A-13507

```

Query Match

```

Best Local Similarity 10.8%; Score 185.6; DB 7; Length 443;
Matches 242; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```

```

Oy 950 AGCTGAGTCCCTGCTACTAGCAAGAAAGTTATGCTGCGAGCATGAGGCTAGAG 1009
Db 98 AACAGTGTACATTTGGGAGCAAGAAAGCTGCTGTTGAACAGGCTGATCCAGAAA 157
Oy 1010 CCACCTACTCTTCAGAAAGCCCAATTTCTATCTCAGAGAGTCAGCCCAATGGCGCT 1069
Db 158 TCGGCACTCTCTCAAAAAGCCAGATTTCTTCAATCTCAGAGGCTCAACCAATGGCAT 217
Oy 1070 TGAACAGTATGCTGACCGGATGAGAGATGAGATGAGATGAGATGAGATGAGAT 1129
Db 218 GAGCAGAGTTTCTCAGATCGTATGATGATGATGATGATGATGATGATGATGAT 277
Oy 1130 TGAAGATCGCCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1189
Db 278 TGAAGATGAGAAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 337
Oy 1190 GCATCTTGGAACTCGTTTGTAGAAAAGAGGTTATAGCAGATGTCATATCTTGT 1249
Db 338 GCATATGTGAAATTCATTTCTGCGAAACAAAGGGTACTAGCGGATGCGCATATTCCTG 397
Oy 1250 GGCATGTGAAGCATTTTTCAGATTTTACGAGAAAGA 1285
Db 398 GGCATGCGAAGCATTTTCTCGATTTTTCATGACAGAGA 433

```

RESULT 10

```

US-10-389-048-21629
; Sequence 21629, Application US/10389048
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilka
; APPLICANT: Shenk, Michael A.
; TITLE OF INVENTION: Polynucleotides, Materials Incorporating

```

```

; TITLE OF INVENTION: Them, and Methods for Using Them
; FILE REFERENCE: 11000.1041u1c1
; CURRENT APPLICATION NUMBER: US/10/389,048
; CURRENT FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 25129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21629
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
;
Db 221 TTGAACAGAAATGTGTAGCAAGAAATGTGCGGAAATTCCTCACCGAGAAAGTATTC 280
Db 64 TTGGGTTCAATGTGCTCCATCAAGTTCTTGTGAACATTTTGTGATGAAGATGCAACTGC 123
Oy 281 AACGATGAGATCTCTTGTATATATTTGTAACCTGTTGCACTATATATACATCTTCACT 340
Db 124 AGCAGAGAGAAATCTCGTGTGATTTGCAAGCTGTTGAGCTATACACATCTTCACTACG 183
Oy 341 TCGCTCTGAGCAACCCATGCTTCTTCCAGATGCTGTAACATAAATTTGAGCAAA 400
Db 184 CCGTCCATACATATATCTTCTTCTCCGAGATGTTGCGTTTCAAAATACAAAGCAAG 243
Oy 401 GCGCAAAAGAAAGTCAAGATCTA--CTGGAGTGTAGTTTCACTATTAAGATGTAA 457
Db 244 AGCTAAGGTTAATTAAGAGGTTGAGGCTGAGATGATGATGATGATGATGATGATGAT 303
Oy 458 TTAACATTAAGAAAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 517
Db 304 CAACGTGATCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 363
Oy 518 ATGTGATGCTTCAAGGGGCTGCATTTCAATTTGAAATTCATCTCATGATTTTGAAT 577
Db 364 ATGTCAAGCTTAAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAG 423
Oy 578 TGAGTCAAGCTTTTCAAGAAATACAGACAGTAAATGTTCTGTAATTAATTCCT 637
Db 424 CGAGTTTGGTATTTGAAGATGATCAAGCAGATGATGATGATGATGATGATGATGAT 483
Oy 638 CATATTTGAGC--AAGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 695
Db 484 GAGATCCGAGGTTCTGACTGAAAGCAACCTAATTTGAGCATTTCTTCTGCGCA 543
Oy 696 AAC 699
Db 544 AAC 547

```

RESULT 11

```

US-10-437-963-91042
; Sequence 91042 Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David R.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 91042
; LENGTH: 2700

```



```

: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 40645
: LENGTH: 322
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(322)
: OTHER INFORMATION: unsure at all n locations
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_136701C.1
: US-10-424-599-40645

```

```

Query Match 7.3%; Score 125.6; DB 10; Length 322;
Best Local Similarity 61.2%; Pred. No. 6.3e-21;
Matches 197; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```

```

OY 1062 ATGGCGCTTGACCAATTAATGCTGACCGGATAGCGAGATGAAGTGAATGAGATGTT 1121
DB 322 ATGGCANTAGAAATTAATGCTGACCGGATAGCGAGATGAAGTGAATGAGATGTT 263
OY 1122 GCAGATTTTGAAGATCGCCAGATGCTTATGATGATGATGATGATGATGATGATG 1181
DB 262 GAGCATGTTGAAGATCGCCAGATGCTTATGATGATGATGATGATGATGATGATG 203
OY 1182 CAATTCATCATCTTTGGAACCTGTTGTAGAAAACAAAGGTTATAGCATGATGAT 1241
DB 202 TAGTTCATCATCTTTGGAACCTGTTGTAGAAAACAAAGGTTATAGCATGATGAT 143
OY 1242 ATCTCTTGCGCATGTAAGATTTTCAAGATTTTACGAAAGAGTTGCAACCGTTACTCA 1301
DB 142 GTTGTGTTGGCTTGTGAGCATTTTCCAGTTTATGAGAAAAGAGCTGATCTATCTCT 83
OY 1302 TCACCTCTGCTGTTGAGATGTTTGTATTAACATGAGAACCATGAGCTTGCGAC 1361
DB 82 GCTTATTTTGGCGTGTGAGGATTTTATGATCAAACTTGGAAATTTATGCTCTTTCAC 23
OY 1362 TCAGCCACCATCAACAACACTGCA 1383
DB 22 GCGCGTACTGATGATTAATGTA 1

```

```

RESULT 15
US-09-837-604B-6414
: Sequence 6414, Application US/09837604B
: GENERAL INFORMATION:
: APPLICANT: BYrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Shukla, Hridayabhiranjan
: APPLICANT: De La Pena, Robert C.
: APPLICANT: Bougri, Oleg
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: FILE REFERENCE: 38-21(51892)B
: CURRENT APPLICATION NUMBER: US/09/837,604B
: CURRENT FILING DATE: 2001-04-18
: PRIOR APPLICATION NUMBER: US 60/197,872
: PRIOR FILING DATE: 2000-04-19
: NUMBER OF SEQ ID NOS: 81288
: SEQ ID NO 6414
: LENGTH: 453
: TYPE: DNA
: ORGANISM: Oryza sativa nipponbare
: FEATURE:
: OTHER INFORMATION: Clone ID: LIB3431-030-PI-K2-G11
: US-09-837-604B-6414

```

```

Query Match 6.0%; Score 104; DB 5; Length 453;
Best Local Similarity 66.5%; Pred. No. 1.6e-15;
Matches 149; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
OY 424 CTGGATGATGATTCTCAACTATAGCATTTGTAATACACATTACAGAAAACGAAAGTTA 483

```

```

DB 173 CTGGCAATGCTCTTTTAACTACAAATACATTAATATGCAAAAAACCGAATCA 232
OY 484 GGGAGGATTTCTTCTGTCATTTTGTCTATGCTATGCTGATGCTCAAGGGGCTGCAAT 543
DB 233 CTGAAGATTTTCTTGGCCATTTTGTCTGTTGATACATGTGCGAGCTTTAAGGCTTAGAT 292
OY 544 TTCAATTTGAATTCATTCATGATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTG 603
DB 293 GTCACTTAACGATTCGATGACCTTTTCCATTATAGTTTGGATATCTGAAGAGTCC 352
OY 604 AGACAGTTAATGTTTCTGTTAAACTTAATTCCTTCATATTGAG 647
DB 353 AGGCTGTTAATGTTATGCTGAAAGACTGATTTCTTGGAAGAGAG 396

```

Search completed: June 19, 2003, 23:39:28
Job time : 1299 secs


```

XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-GB00248.
XX
XX 28-JAN-1999; 99GB-0001927.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Dean C, Gendall A;
XX WPI: 2000-499333/44.
XX N-PSDB; AAA47751.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
XX with altered vernalization response, flowering time, leaf size and/or
XX shape or shade avoidance response for maximized reproductive success
XX
XX Claim 49; Fig 6; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX plant encode polypeptides which are capable of affecting one or more
XX vernalization responses such as, flowering time, leaf size and/or
XX shape or the shade avoidance response of a plant into which the
XX nucleic acid is introduced. Introducing such sequences into plants
XX to alter these characteristics maximises the reproductive success of
XX the plant.
XX
XX Sequence 445 AA;
SQ
Query Match 100.0%; Score 2378; DB 21; Length 445;
Best Local Similarity 100.0%; Pred. No. 2.4e-234;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSPFLPCLNRYKIGAKRR 60
DB 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSPFLPCLNRYKIGAKRR 60
QY 61 KSRSTGMVYVNYKDCNNITLQKTEVREDSCSPFCMCGSKGLQFHLNSHDLFEFEFKL 120
DB 61 KSRSTGMVYVNYKDCNNITLQKTEVREDSCSPFCMCGSKGLQFHLNSHDLFEFEFKL 120
QY 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
QY 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
QY 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEKTSNIPPAIHSSLDAGAKYILTSE 240
DB 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEKTSNIPPAIHSSLDAGAKYILTSE 240
QY 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRDSEDDVDVADF 300
DB 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRDSEDDVDVADF 300
QY 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
DB 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
QY 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
DB 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
QY 361 MCWRLEFLIKLMNHGLVDSATINNCTILLECRNSSDTTNNNNNSVDSPDSNTNNNNIY 420
DB 361 MCWRLEFLIKLMNHGLVDSATINNCTILLECRNSSDTTNNNNNSVDSPDSNTNNNNIY 420
QY 421 DHPNDJNNKNNVNNKNNNSRDKYIK 445
DB 421 DHPNDJNNKNNVNNKNNNSRDKYIK 445

```

```

DT 16-NOV-2000 (first entry)
XX
XX VRN2 polypeptide.
DE
XX
XX Vernalization gene: VRN2; plant characteristic: flowering time;
XX leaf size; leaf shape; shade avoidance response; reproduction;
XX breeding; pollination; cultivation.
XX Arabidopsis thaliana (Columbia).
XX WO200044918-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-GB00248.
XX
XX 28-JAN-1999; 99GB-0001927.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Dean C, Gendall A;
XX WPI: 2000-499333/44.
XX N-PSDB; AAA47753.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
XX with altered vernalization response, flowering time, leaf size and/or
XX shape or shade avoidance response for maximized reproductive success
XX
XX Claim 47; Page 71-72; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX plant encode polypeptides which are capable of affecting one or more
XX vernalization responses such as, flowering time, leaf size and/or
XX shape or the shade avoidance response of a plant into which the
XX nucleic acid is introduced. Introducing such sequences into plants
XX to alter these characteristics maximises the reproductive success of
XX the plant.
XX
XX Sequence 440 AA;
SQ
Query Match 96.1%; Score 2285; DB 21; Length 440;
Best Local Similarity 96.8%; Pred. No. 7.8e-225;
Matches 428; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
QY 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSPFLPCLNRYKIGAKRR 60
DB 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSPFLPCLNRYKIGAKRR 60
QY 61 KSRSTGMVYVNYKDCNNITLQKTEVREDSCSPFCMCGSKGLQFHLNSHDLFEFEFKL 120
DB 61 KSRSTGMVYVNYKDCNNITLQKTEVREDSCSPFCMCGSKGLQFHLNSHDLFEFEFKL 120
QY 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
QY 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 121 FEETQYVAVSVKINSFIPEEGSDDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
QY 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEKTSNIPPAIHSSLDAGAKYILTSE 240
DB 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEKTSNIPPAIHSSLDAGAKYILTSE 240
QY 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRDSEDDVDVADF 300
DB 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRDSEDDVDVADF 300
QY 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
DB 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
QY 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
DB 301 EDROMLDDFVDVNRKDKQFMHLMNSFVRKQVIAQGHISMACEASRFEYKEKELHRYSLF 360
QY 361 MCWRLEFLIKLMNHGLVDSATINNCTILLECRNSSDTTNNNNNSVDSPDSNTNNNNIY 420
DB 361 MCWRLEFLIKLMNHGLVDSATINNCTILLECRNSSDTTNNNNNSVDSPDSNTNNNNIY 420

```

OY 421 DHPNDINNNKNNVDKNNRSRK 442
||||| |||||||||
Db 419 DHPNDIKNNKNVDKNNRSRK 440

RESULT 3
AAC54168
ID AAC54168 standard; Protein; 498 AA.
XX
AC AAC54168;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69039.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156539.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158363.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161923.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.9%; Score 1923; DB 21; Length 498;
Best Local Similarity 89.4%; Fred. No. 1,le-187;
Matches 370; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

OY 43 PSLIPRLNKKIKAKKRRKSRSTGAVVFNKYKDCNNTLOKTEVREDSCPCFSCMLGSEFKG 102
DB 108 PSLIPRLNKKIKAKKRRKSRSTGAVVFNKYKDCNNTLOKTEVREDSCPCFSCMLGSEFKG 167
OY 103 LQHLNSSHDLFEFEKLFEEYQTVNVSVKLSFTFEERGSDDKFEFESLCSKPKRKRO 162
DB 168 LQHLNSSHDLFEFEKLFEEYQTVNVSVKLSFTFEERGSDDKFEFESLCSKPKRKRO 204

OY 163 RGRNNTRELKACFLPLDSPSLTNGTNGTITLLDNGRSLGYPEATELIGOFEMTNIIP 222
DB 205 RGRNNTRELKACFLPLDSPSLTNGTNGTITLLDNGRSLGYPEATELIGOFEMTNIIP 264
OY 223 ATAHSSLDGAKAVILTSAAVPAKTKRTISAERSEAR -----SHILOKROFYH 271
DB 265 ATAHSSLDGAKAVILTSAAVPAKTKRTISAERSEARSTAILTEKSSSHLOKROFYH 324
OY 272 SHRVOPMALEOVMSRDEDEVDVADFEEDROMLDFVYVNDKDEQFPHLNNSTYRKOR 331
DB 325 SHRVOPMALEOVMSRDEDEVDVADFEEDROMLDFVYVNDKDEQFPHLNNSTYRKOR 384
OY 332 VIADGHISMACEAFSEFEKELHRYSSILFCWRLFLIKIMHGLVDSATINNCNTILENC 391
DB 385 VIADGHISMACEAFSEFEKELHRYSSILFCWRLFLIKIMHGLVDSATINNCNTILENC 444
OY 392 RNSDPTTTNNNSVDRPSDNTNNNIYDHPNDINNNKNNVNDKNNNSDKYIK 445
DB 445 RNSDPTTTNNNSVDRPSDNTNNNIYDHPNDINNNKNNVNDKNNNSDKYIK 498
RESULT 4
AAC54169
ID AAC54169 standard; Protein; 367 AA.
XX
AC MAG54169;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69040.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132487.
PR 05-MAY-1999; 99US-0132488.
PR 06-MAY-1999; 99US-0132489.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136621.
PR 27-MAY-1999; 99US-0136692.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147703.
PR 06-AUG-1999; 99US-0147716.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.


```

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 03-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.

```

```

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.8%; Score 1635; DB 21; Length 339;
Best Local Similarity 87.8%; Pred. No. 1.8e-158;
Matches 318; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

OY 95 MLCGFKGLQFLHNSHDLFEFEKLFEEYOTVNVSVKLSFIIEEGSDDDKEPFSIC 154
DB 1 MLCGFKGLQFLHNSHDLFEFEKLFEEYOTVNVSVKLSFIIEEGSDDDKEPFSIC 37
OY 155 SKPRRRRGGRNNRRRLKVCFLPLDPSLTNGTENGITLLDNGRGIGYEATLGAQF 214
DB 38 SKPRRRRGGRNNRRRLKVCFLPLDPSLTNGTENGITLLDNGRGIGYEATLGAQF 97
OY 215 EMTSNIPPAIHSSIDAGAKYILTSEAVPATKTKLSAESEAR-----SHLL 263
DB 98 EMTSNIPPAIHSSIDAGAKYILTSEAVPATKTKLSAESEARSTAILTEKISSILL 157
OY 264 LÖKROFYSHRYVQPMALFQVMSDRDSEDDVDVADFEDROMLDPVDVNDKQFPHLM 323
DB 158 LÖKROFYSHRYVQPMALFQVMSDRDSEDDVDVADFEDROMLDPVDVNDKQFPHLM 217
OY 324 NSFVRKQRYVADGHISWACEAFSRYEKELHRYSSLEWCWRLFILKLMNHGLVDSATINN 383
DB 218 NSFVRKQRYVADGHISWACEAFSRYEKELHRYSSLEWCWRLFILKLMNHGLVDSATINN 277
OY 384 CNTILENCRNSSDTTNNNSVDPSDNTNNNNIYDHPMDINNNKNNVDKNNNSRDYV 443
DB 278 CNTILENCRNSSDTTNNNSVDPSDNTNNNNIYDHPMDINNNKNNVDKNNNSRDYV 337
OY 444 IK 445
DB 338 IK 339

RESULT 6
AAV53932
ID AAV53932 standard; Protein; 611 AA.
XX AAV53932;
AC
AC AAV53932;
XX
XX 13-MAR-2000 (first entry)
DT
DE A MPC1 protein having flowering regulating activity.
XX
XX
XX Flowering regulating activity; MPC1; flowering; germination;
KW super early flowering mutation; altered flowering time;
KW flowering regulating gene; food crop; vegetable; flowering inhibition;
KW productivity.
XX
XX Arabidopsis thaliana.
OS
XX

```

Key	Location/Qualifiers
Domain	306..327
	/note="zinc finger domain"
EP967278-A2.	
29-DEC-1999.	
28-JUN-1999.	99EP-0305077.
26-JUN-1998.	98JP-0180065.
24-JUN-1999.	99JP-0179043.
(MITA) MITSUI CHEM INC.	
Yoshida N, Kato Y, Takahashi S, Yanaai Y, Hiratsuka J, Miwa T;	
N-PSDB: AA236947, AA236948.	
Novel DNA used to produce transgenic plants with altered floral regulation which can have increased crop yields -	
Claim 2, Page 12-15; 53pp; English.	
The present sequence represents a protein having a flowering regulating activity, which is designated MPC1. The genomic sequence is given in AA236948. A rice MPC1 is also disclosed in the specification. The rice and Arabidopsis cDNAs show significant homology with each other. A naturally occurring mutation of the MPC1 gene eliminates normal flowering regulating ability of plants, and leads to flowering immediately after germination (super early flowering mutation), and leads to a truncated protein complementing amino acids 1-540 of the present sequence. The MPC1 polynucleotide sequence can be used to produce plants with altered flowering times in comparison with wild type plants, by enhancing or inhibiting the expression of the flowering regulating gene. Antisense polynucleotides can be used to reproduce the effects of the mutated MPC1 gene. This alteration can be used to increase the yield of food crops. Flowering inhibition of vegetables increases their productivity.	
Sequence 611 AA:	
Query Match	38.8%; Score 923; DB 21; Length 611;
Best Local Similarity	34.8%; Pred. No. 3.5e-85;
Matches 216; Conservative 52; Mismatches 108; Indels 244; Gaps 9;	
1 MCRONCRASSEFEYISTDENLLIYCKPRLNIFHLRSLGNPSFLPCLINRYIGAKRRR 60	
1 MCHEDSRLLTISEEBELAAESLIAYCKRVELYNILIORRAIRNPLFLQRCILHYKIEAKHRR 60	
61 -----	60
61 RIOMTVLSGAIDAGVOTOKLEPLYILLARLVSPKPYAEYSAAVYRFRACITLGLGVDC 120	
61 -----	60
61 VSOAQAANFLLPDMNRLLAEAKSGSLAILFISFAGQANSPGIDSGRIHSGNIGHCIMSK 180	
61 -----	64
61 IFLQSLIYASWQSPMMDDGQRVDIVSLVEMQPCFIRKLSSEKCVSIQVSPNPLTSSSP 240	
65 -----	65
65 OGVQVYISAEVGTSEKSPYSSFSYNISSSLILQILRLTGAVVNYRYNNKIQKTEY 300	
85 RRDSCSPFCMSLGCSEFKGIQPHLSSHDLEFEFKLFEEOVNVSVKLNSLTFEEESGD 144	
301 TEDFSCPCFLYKCAAFKGLRHLSTHDLNLFEEWYIEEQAVNSLKTETMSKYNEDD 360	
145 -DDKREPSLCSKP-RKRROGGRNNRRLKVCFLPLDPSLLNG-----TENGITLLN 196	

Dd		361	VDRKQCTFEFFSSKKFRRRRQKQVASSNQ-----GPHLGCGEVLKTDKDAHSVR	411
Oy		197	DGR---GLGYPRATLGAQFEMTSNIP-----PAIASHSLDAGAKVIITSNAVYPATKR	249
Dd		412	EKNRIPEPKHY-----BRIGASGQRVPFGSPISPADVOGCDDPYDVOSTAGSTMLQFAKTR	467
Oy		250	KLSAEFSSEARSHLLLOKROFYTHSRHYPMALEQWMSDPDSDEYVDVADPEDROMIDDF	309
Dd		468	KTIESSDLRNNSLLOKROFHFSSHRAQPMALLESQVLDSDSDSEYDDVADPEDERKLDDF	527
Oy		310	VDVNCKEQPMLMNSFVRKORVIADGHISMACEASFSEYKEELHRYSLSFWCWRLEFLIK	369
Dd		528	VDVTKEKQMHHMMNSFVRKORVLDGHITPMACEAFSRLHGPIVKRPPLILWCVRVMVK	587
Oy		370	LWNHGLVDSATINCNNTILE	389
Dd		588	LWNHGLIDARTIMNCNTFLE	607
RESULT 7				
AAV53933				
ID	AAV53933	standard; Protein; 604 AA.		
XX	AAV53933;			
AC				
XX				
DT	13-MAR-2000	(first entry)		
XX				
DE	An Os-MPC1 protein having flowering regulating activity.			
XX				
KW	Flowering regulating activity; MPC1; flowering; germination;			
KM	super early flowering mutation; altered flowering time;			
KW	flowering regulating gene; food crop; vegetable; flowering inhibition;			
KM	productivity.			
OS				
XX	Oryza sativa.			
FH				
XX	Key	Location/Qualifiers		
FT	Domain	310..335		
FT		/note="zinc finger domain"		
PN	EP967278-A2.			
PD				
XX	29-DEC-1999.			
EF	28-JUN-1999;	99EP-0305077.		
XX				
PR	26-JUN-1998;	98JP-0180065.		
PR	24-JUN-1999;	99JP-0179043.		
PA	(MITA) MITSUI CHEM INC.			
XX				
PI	Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;			
XX				
DR	WPI; 2000-064612/06.			
DR	N-PSDB; AAZ36953.			
PT				
PT	Novel DNA used to produce transgenic plants with altered floral			
XX	regulation which can have increased crop yields -			
PS	Claim 3; Page 36-39; 53pp; English.			

The present sequence represents a protein having a flowering regulating activity, which is designated Os-MPC1. An Arabidopsis cDNA is also disclosed in the specification. The rice and Arabidopsis cDNAs show significant homology with each other. A naturally occurring mutation of the MPC1 gene eliminates normal flowering regulating ability of plants and leads to flowering immediately after germination (super early flowering mutation). The MPC1 polynucleotide sequence can be used to produce plants with altered flowering times in comparison with wild type plants, by enhancing or inhibiting the expression of the flowering regulating gene. Antisense polynucleotides can be used to reproduce the effects of the mutated MPC1 gene. This alteration can be used to increase the yield of food crops. Flowering inhibition

CC of vegetables increases their productivity.
XX Sequence 604 AA;

Query Match 38.1%; Score 906.5; DB 21; Length 604;
Best Local Similarity 33.3%; Pred. No. 1.7e-83;
Matches 209; Conservative 59; Mismatches 109; Indels 251; Gaps 7;

QY 1 MCRONCRASSPEEYSTDENLLIYCKPRLNFIHLRSGNSFPLPCLNTYTGAKRR 60
||| ||| :||| :||| :||| ||| ||| :||| :||| ||| ||| ||| :|||
DB 1 MCRHOPRRLSPDEOLAESEFALYCKPVELNIIORRSIKNPAFLQCLIKIARRR 60
QY 61 KSRST----- 65
|||
DB 61 RSLITISLSGTNKLRAQNIPLVYLARPTNNVSLSGHSPIYRSRACLLTSFHFQ 120
66 ----- 65
QY 121 KDYTEATFVLPVKNIATSRACSLNIIISCGRAEQTFDDNNGSNHVEGSTLQLEK 180
QY 66 ----- 65
DB 181 FMGKIPIDILASSLNCVSLSGHTVEMSTVEMTPSFLEPKFLEDSCLTFCGQVDA 240
QY 66 ----- -GMVFENYKDCNNLTQ 80
DB 241 GSRQLQVSIQAQAGAKMSESPYSYVSYNDVPPSSLTHIRLRSGNVLENYKYNNT 300
QY 81 KTEVREDCSCPFCSMLGSGFKGLQFHLNSHDLFEFEKLFEEYQTVNVSVKLSFIF 139
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 KTEVREDCSCPFCLVPCGSGFKGLCHLNASHDLFHYEFWISSECAVNVSLKIDSWR 360
QY 140 -EESGDDDDKEEPPSLCSKPRKR-RQRGNNRRLKLVCLPLDSP-SLNGTPNGTLL 196
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 LAEGV-DPRHQTFSTYSRKRKRKRVISSTKTHVPHIVDSGSPDAQAGSDDYQRE 419
QY 197 DGNRGLGYEATELACQFEMTSNIPALHSSLDAGAKV---LITSEAVVPATKTKLS 253
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 NGS-----SVAAVSDPANSLHGSNLSA.PTVLQFCKTKLS 456
QY 254 ERSEARSHLLQKQFYHSHRVPALBOYMSRDESEDDVADFEQRMDDVDVNA 313
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 457 ERADPPNRQLQKQFHSRAQPMAMSKVFSRDRSEDEDDIADFEERMLDDVDV 516
QY 314 KDEKQFMHLMNSFVRQRYIADGHISMACEAFGRFEKLEHRSLSLFCWRLFLIK 373
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 517 KDEKLIMHNMNSFVRQRYIADGHIPACEAFSGFHGQELVQNPALIMCKRFPV 576
QY 374 GLVDSATINNCTIENCRRNSDPTTNN 401
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 577 SLIDARAMNACNTILEGYLNGSSDPKK 604

RESULT 8
AAB00062
ID AAB00062 standard; Protein; 107 AA.
XX
AC AAB00062;
XX
DT 16-NOV-2000 (first entry)
XX
DE VRN2 polypeptide generated from aberrantly spliced VRN2 nucleic acid.
XX
KM Vernalization gene; VRN2; plant characteristic; flowering time;
KM leaf size; leaf shape; shade avoidance response; reproduction;
KM breeding; pollination; cultivation.
XX
OS Arabidopsis thaliana (Columbia).
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX

PF 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI: 2000-499333/44.
DR N-PSDB: AAA47755.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Claim 47; Page 75; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This polypeptide resulted from an aberrant splice in the
CC VRN2 gene.
XX
SQ Sequence 107 AA;

Query Match 23.1%; Score 549; DB 21; Length 107;
Best Local Similarity 99.0%; Pred. No. 4.9e-48;
Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCRONCRASSPEEYSTDENLLIYCKPRLNFIHLRSGNSFPLPCLNTYTGAKRR 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MCRONCRASSPEEYSTDENLLIYCKPRLNFIHLRSGNSFPLPCLNTYTGAKRR 60
QY 61 KSRSTGMVFENYKDCNNLTQKTEVREDCSCPFCSMLGSGFK 101
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 KSRSTGMVFENYKDCNNLTQKTEVREDCSCPFCSMLGSGFK 101

RESULT 9
AAB01674
ID AAB01674 standard; protein; 813 AA.
XX
AC AAB01674;
XX
DT 01-SEP-2000 (first entry)
XX
DE FIS2 protein sequence.
XX
KM Seed; development; FIS; endosperm; autonomous embryogenesis;
KM transgenic plant; seedless fruit; parthenocarpic; citrus fruit;
KM stone fruit.
XX
OS Arabidopsis thaliana.
XX
FH Key location/Qualifiers
FT Misc-difference 1..120
FT /note="The first 120 residues are illegible in the
FT specification"
XX
PN WO200016609-A1.
XX
PD 30-MAR-2000.
XX
PF 21-SEP-1999; 99WO-AU00805.
XX
PR 21-SEP-1998; 98US-0101184.
PR 22-SEP-1998; 98AU-0006061.
PR 22-SEP-1998; 98AU-0006062.
PR 22-SEP-1998; 98AU-0006063.
PR 01-JUL-1999; 99AU-0001345.

PR 01-JUL-1999; 99AU-0001346.
 XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
 PA Blodiadeu P, Chaudhury AM, Dennis ES, Koltunow AMG, Luo M;
 PI Peacock WJ;
 PI WPI: 2000-283392/24.
 DR
 XX Induction of seed development in plants in the absence of fertilization
 PT by inhibiting or preventing the expression of a negative regulator of
 PT seed formation for production of seedless or soft-seeded fruit
 XX
 XX Claim 2; Fig 14; 207pp; English.
 XX
 CC The present invention relates to a method of inducing the development of
 CC seeds in a plant, comprising inhibiting, interrupting or reducing the
 CC expression of a negative regulator of seed formation in one or more
 CC female reproductive cells, tissues, or organs of the plant or a
 CC progenitor cell, tissue or organ. The negative regulator is a
 CC polypeptide. The present sequence represents the FIS2 protein sequence.
 CC The FIS family of genes are known to be capable of regulating autonomous
 CC endosperm development and/or autonomous embryogenesis. In the invention
 CC the reduced expression of the negative regulator is achieved by the
 CC introduction of a transgene which comprises a FIS genetic sequence, which
 CC may inhibit FIS activity. Plants produced using the method of where the
 CC invention produce parthenocarpic fruit or soft-seeded fruit
 CC fruit are made parthenocarpic or have soft seed by a process comprising
 CC expressing the introduced nucleic acid molecule in a tissue or organ of
 CC the fruit. The plant produces seed independent of fertilization. The
 CC isolated FIS nucleic acid molecules are used in the production of an
 CC antisense molecule, a ribozyme, a co-suppression molecule, a
 CC gene-targeting molecule, a gene-silencing molecule and a
 CC dominant-negative sense molecule where the member is used for the
 CC production of a transformed plant. The transformed plant is apomictic or
 CC produces soft-seeded or parthenocarpic fruit. Production of soft-seeded
 CC fruit has large economic value, since it makes the fruit more desirable
 CC to customers. Examples include stone fruits such as apricots and peaches,
 CC citrus fruits such as oranges, lemons, grapefruit and mandarins and other
 CC fruits such as grapes, apples, pears and berries. The plants
 CC which undergo autonomous seed formation do not require fertilization to
 CC reproduce, and may express desirable characteristics stably between
 CC generations. Antibodies produced to the FIS polypeptides can be used to
 CC detect the peptides of the invention and can be used in an enzyme linked
 CC immunosorbant assay (ELISA), radioimmunoassay or histochemical tests.
 XX
 XX Sequence 813 AA;
 S0
 Query Match 21.5%; Score 510.5; DB 21; Length 813;
 Best Local Similarity 22.3%; Pred. No. 9e-43;
 Matches 154; Conservative 53; Mismatches 121; Indels 363; Gaps 8;
 QY 81 KTEVREDSCPCFCSCFSGFKGLQFLHNSHDLFFEEKFLPEEYQ--TVNVSVLNLSIF 138
 DB 125 KAEVVENFSCPFCLICGCGHEGLQHLKSHDAFFERYRAEKDQGPVDVSVASDRIKF 184
 QY 139 E--EGSDDDKPEFSLGSKPR-KRRORGRNNTRLVWCFLPDSLTGCTGTLL 195
 DB 185 GVLKDVGNQPLSLPLFGSKNNORGRDSSNNVKLVLMELDDLPDGTENDSTH 244
 QY 196 NDCN----- 199
 DB 245 NDNVSSPPRAHSEKISDILTTTQLAIESSEPPVPHVNGVNSPPRAHSSAKNEST 304
 QY 200 ----- 199
 DB 305 HVNDDDDVSSPPRAHSEKNESTHVNEDNISPPRAHSSKKNSTHNDVDSFPPTRS 364
 QY 200 -----RG----- 201
 DB 365 SKTSDILTTTQPAIYSEPPVRRGSRKROLYAKRYKARTOPAIASSEPKVLHVND 424
 QY 202 --LGYPEATELAGOEFMTSNIPPAIHSS----- 228

DB 425 NVSSPEAHSLKASDILTTTQPAIASESEPKVPHVNDVNSSTPRAHSSKKNKSTPKNV 484
 QY 229 ----- 228
 DB 485 DVNSPPKTRSSKKTSDILTTTQPAIASESEPKVPHVNDVNSSTPRAHSSKKNKSTRN 544
 QY 229 -----LDGAKVILTSVAVP----- 244
 DB 545 DNIPSPPKTRSSKKTSLNLRTPAIASESPKVPVHNDKVSSTPRAHSSKKNKTHK 604
 QY 245 -----ATRKLSAERSPARSHLLQKR 267
 DB 605 DNASLPPKTRSSKKTSDILATTPAKAEPSEPKVTRSVSRKELHARECAKRLERLGR 664
 QY 268 OFYHSRRYQPMALREQVMSDRDSEDEVDVADVEDRMDPVDVNDKDEKOFWHLNPFV 327
 DB 665 OFYHSOTQMPFTEQVMSDESENEETDVALDISERLRLVGVSEKERYWYLMNIFV 724
 QY 328 KROYIADGHTSMACEAREPKEKELHRYSLFWCRLFLIKLNNIGVDSATINNCTI 387
 DB 725 KROYIADGHTSMACEAREPKEKELHRYSLFWCRLFLIKLNNIGVDSATINNCTI 784
 QY 388 LENCNNSDTTNNNSVDRPSDSTNNNN 418
 DB 785 L-----LSNDEAGOTTSGSAANANN 805
 RESULT 10
 AAB00065
 ID AAB00065 standard; Protein: 186 AA.
 XX
 AC AAB00065;
 XX
 DT 16-NOV-2000 (first entry)
 DE At HYP 2245035 (ATFCA7_4) (modified).
 KW Vernalization gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation.
 XX
 OS Unspecified.
 XX
 PN WO200044918-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-GB00248.
 XX
 PR 28-JAN-1999; 99GB-0001927.
 PA (PLAN-) PLANT BIOSCIENCE LTD.
 PI Dean C, Gendall A;
 PI WPI: 2000-49933/44.
 DR N-PSDB: AAA47758..
 XX
 PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 XX
 PS Disclosure: Page 77; 105pp; English.
 XX
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant. This cDNA was generated from an aberrantly spliced VRN2
 CC gene.

Fri Jun 20 09:03:39 2003

us-09-890-220-2.rag

Page 12

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139152.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142300.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145518.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145591.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152370.
PR 11-SEP-1999; 99US-0153708.
PR 12-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 327; DB 21; Length 63;
Best Local Similarity 96.8%; Pred. No. 1,2e-25;
Matches 61; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 383 NCNTILENCRSSSTTTTNNNNNSYDRPSDSTNNNTVYDHPNDINNNKNNYDNKDNNSRDK 442
Db 1 NCNTILENWRNSSDTTTNNNNNSYDRPSDSTNNNTVSHPNIDINNNKNNYDNKDNNSRDK 60

OY 443 VIK 445
Db 61 VIK 63

RESULT 13
AAG61332
ID AAG61332 standard; Protein: 145 AA.

AC AAG61332;
DY 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 79541.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX

PE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153078.
PR 13-SEP-1999; 99US-0153750.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 324; DB 21; Length 145;
Best Local Similarity 98.3%; Pred. No. 8.3e-25;
Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 43 PSELPRLNRYKIGAKRRKRSRSTGMVFNKDCNNTLQKTEVEDSCSPFCSMTCGSPK 101
|||||
Db 81 PSELPRLNRYKIGAKRRKRSRSTGMVFNKDCNNTLQKTEVEDSCSPFCSMTCGSPK 139
RESULT 14
AAC61331
ID AAC61331 standard; Protein: 174 AA.
XX
AC AAC61331;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79540.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
PN EP103405-A2.
PD
XX
XX 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132407.
PR 06-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132853.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137529.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138064.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.


```

AAB00063
ID AAB00063 standard; Protein; 154 AA.
XX
AC AAB00063;
XX
DT 16-NOV-2000 (first entry)
XX
DE C72616 Expressed sequence tag (modified).
XX
XX Vernalization gene: VRN2; plant characteristic: flowering time;
KM leaf size; leaf shape; shade avoidance response; reproduction;
KM breeding; pollination; cultivation.
XX
OS Unspecified.
XX
PN WO200044918-A1.
XX
PD 03-AUG-2000.
XX
FE 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI: 2000-499333/44.
XX
DR N-PSDB; AAA47756.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Disclosure; Page 76; 105pp; English.
XX
CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA was generated from an aberrantly spliced VRN2
CC gene.
XX
SQ Sequence 154 AA;
XX
Query Match 13.1%; Score 311; DB 21; Length 154;
Best Local Similarity 62.1%; Pred. No. 2e-23;
Matches 64; Conservative 14; Mismatches 21; Indels 4; Caps 2;
OY 223 ATAHSSLDAGAKV--ILTSEAVVPATYTRKLSAERSEARSHLLQKRFYHSHRVOPMA 279
Db 53 SYAHASVDPANSLHGSNLSAPTVLQFGTKRLSYERADPRNRQLQKRFHSHRAQPM 112
OY 280 LEOVMSDRDSEDEVDDVADEFEDROMLDDFVDVANKDEKQFMHL 322
Db 113 LGAVFSDRDSDEVDVDDIADFEEDROMLDDFVDVYKDF-LIMHM 154

```

Search completed: June 13, 2003, 15:35:41
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 13, 2003, 15:37:19 ; Search time 26 Seconds

(without alignments)
503.585 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRACKSPPEVISTDE.....INNNKNDKNSRDYK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

262574

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126.5	5.3	1584	US-09-457-040B-27	Sequence 27, Appl
2	123	5.2	110	US-08-569-166-34	Sequence 34, Appl
3	112	4.7	361	US-08-415-751-3	Sequence 3, Appl
4	105	4.4	732	US-08-914-999-8	Sequence 8, Appl
5	103	4.3	432	US-08-700-152A-4	Sequence 4, Appl
6	103	4.3	1073	US-09-541-782-6	Sequence 6, Appl
7	103	4.3	1073	US-09-723-820-6	Sequence 6, Appl
8	98.5	4.1	1388	US-09-572-191-2	Sequence 2, Appl
9	98.5	4.1	1388	US-09-723-262-2	Sequence 2, Appl
10	98.5	4.1	1388	US-09-723-219-2	Sequence 2, Appl
11	98	4.1	652	US-08-559-896B-2	Sequence 2, Appl
12	94	4.0	2391	US-08-446-855A-2	Sequence 2, Appl
13	94	4.0	2391	US-09-150-741-2	Sequence 2, Appl
14	92	3.9	286	US-09-384-162-6	Sequence 2, Appl
15	91.5	3.8	689	US-09-177-249-2	Sequence 2, Appl
16	91.5	3.8	689	US-09-061-769A-2	Sequence 2, Appl
17	91.5	3.8	832	US-08-677-734A-12	Sequence 12, Appl
18	91.5	3.8	832	US-09-097-053-12	Sequence 12, Appl
19	91.5	3.8	1786	US-08-973-462-8	Sequence 8, Appl
20	91	3.8	1051	US-08-965-762-14	Sequence 14, Appl
21	91	3.8	1051	US-09-911-927-14	Sequence 14, Appl
22	91	3.8	1051	US-09-911-882-14	Sequence 14, Appl
23	91	3.8	1099	US-09-442-100-2	Sequence 2, Appl
24	90.5	3.8	428	US-09-118-319-5	Sequence 5, Appl
25	90.5	3.8	581	US-08-619-812-6	Sequence 6, Appl
26	89.5	3.8	826	US-07-638-431-2	Sequence 2, Appl
27	89.5	3.8	826	PCT-US92-00018-2	Sequence 2, Appl

28	89	3.7	749	2	US-08-568-459A-6	Sequence 6, Appl
29	89	3.7	749	2	US-08-487-826B-6	Sequence 6, Appl
30	89	3.7	749	4	US-09-210-288-6	Sequence 6, Appl
31	88.5	3.7	432	3	US-09-118-319-2	Sequence 2, Appl
32	88.5	3.7	507	4	US-09-134-001C-3978	Sequence 3978, Ap
33	88.5	3.7	596	4	US-09-752-165-2	Sequence 2, Appl
34	88.5	3.7	911	3	US-09-074-579-4	Sequence 4, Appl
35	88.5	3.7	911	3	US-09-388-774-4	Sequence 4, Appl
36	88	3.7	930	4	US-09-134-001C-5314	Sequence 5314, Ap
37	87.5	3.7	392	3	US-08-523-373-6	Sequence 6, Appl
38	87.5	3.7	1081	2	US-08-843-530B-18	Sequence 18, Appl
39	87.5	3.7	1117	2	US-08-843-530B-33	Sequence 33, Appl
40	87	3.7	620	1	US-08-484-105-4	Sequence 4, Appl
41	87	3.7	620	1	US-08-484-105-4	Sequence 4, Appl
42	87	3.7	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
43	86.5	3.6	770	2	US-08-209-521-13	Sequence 13, Appl
44	86.5	3.6	770	4	US-08-961-810-123	Sequence 123, App
45	86.5	3.6	770	4	US-08-352-902D-123	Sequence 123, App

ALIGNMENTS

```

RESULT 1
US-09-457-040B-27
; Sequence 27, Application US/09457040B
; Patent No. 6387641
;
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: DICDI - Dictyostellium Discoideum
US-09-457-040B-27

Query Match      5.3%; Score 126.5; DB 4; Length 1584;
Best Local Similarity 22.2%; Pred. No. 0.00056;
Matches 77; Conservative 37; Mismatches 100; Indels 133; Gaps 16;

QY 149 EEFSLCKRKRKRGGRNTRRLKVCFLDPSLTNGTGLTLNDGNGLGPEAT 208
   |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
DB 210 EEEFVCRMSKKNYH-GKHVVYRNKRTAFPLDSPDITG---GVRATOPFGGFCY 261
QY 209 ELAOGFEMTSNIPPAIAHSSLDGAKVILTSEAVVPATKTRKLSAERSEARSHLLLOKRO 268
   |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
DB 262 -----FE-----VIIDQLDKGOLSLISLANLEKPTF-----YHGVCMPRS 295
QY 269 F-YHSHRVPALAEQMSDRDSEDEYDDVADFEEDROMIDVDVYVKNDEKQMHLMNSFV 327
   |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
DB 296 YGYH-----NDDGRFRFRNKEE-----PGVNEGSS-----YGSY 324
QY 328 RKQRYIADGHSIMACE-----AFSRFEKELHRRYSLFMCWRLE----- 366
   |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
DB 325 KKGDIITGGC-LSFSTRREIFFTNGMYLGTAFSNVT-----GFYPSVAANEPCISIT 375
QY 367 -----LRLKMHGLVDSATI-----NMCNTIELNCRNSSDPT 398
   |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
DB 376 GYGFPPKFSQVTLMLKN---VNSGSLVLPVNNNNNNNNNNNNNNNNNNNNNNNNNNNN 432
QY 399 TT-----NNNNSVDRPSDSTNNNNNYVDHPNDINKNNKNDKNNNS 439
   |||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.|||.
DB 433 TSTSPSINNEDISSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 479

```

```
? Sequence 34, Application US/08569166
? Patent No. 5830722
? GENERAL INFORMATION:
? APPLICANT: NICOLAS, LUC
? APPLICANT: CHARLES, JEAN-FRANCOIS
? APPLICANT: DELECLUSE, ARMELE
? APPLICANT: BARLOY, FREDERIQUE
? TITLE OF INVENTION: CLOSTRIDIUM BIFERMENTANS DNA FRAGMENT
? TITLE OF INVENTION: BEARING GENES CODING FOR PROTEINS LINKED TO AN
? TITLE OF INVENTION: INSECTICIDAL ACTIVITY
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
? STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
? CITY: ARLINGTON
? STATE: VA
? COUNTRY: USA
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/569,166
? FILING DATE: 05-JUL-1996
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/FR94/00768
? FILING DATE: 24-JUN-1994
? PRIOR APPLICATION NUMBER: FR 93/07795
? FILING DATE: 25-JUN-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: OBLON, NORMAN F.
? REGISTRATION NUMBER: 24,618
? REFERENCE/DOCKET NUMBER: 660-106-0 PCT
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-3000
? TELEFAX: 703-413-2220
? INFORMATION FOR SEQ. ID NO: 34:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 110 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-569-166-34

Query Match 5.2%; Score 123; DB 2; Length 110;
Best Local Similarity 38.7%; Pred. No. 1.8e-05;
Matches 24; Conservative 16; Mismatches 18; Indels 4; Gaps 1;

OY 380 TNNCNTILNCNRSDDTTTNNNNNSVDRPSDSTNTNNNIYDHPNDINNKNNVYDNKDNNS 439
Db 10 TNNNTNNNNNNNNNNNNNNNNNNNA---NNQNTNNNNMGNNSNNNNNNPNNNNHQQNNNN 65

OY 440 RD 441
Db 66 NN 67

RESULT 3
? Sequence 3, Application US/08415751
? Patent No. 5643772
? GENERAL INFORMATION:
? APPLICANT: PETERSEN, CAROLYN
? APPLICANT: LEBCH, JAMES
? APPLICANT: NELSON, RICHARD, C.
? APPLICANT: GUT, JIRI
? TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
```

```
? TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
? TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
? TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
? TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
? NUMBER OF SEQUENCES: 50
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PHILLIPS, MOORE, LEMPIO & FINLEY
? STREET: 385 Sherman Avenue, Suite 6
? CITY: Palo Alto
? STATE: California
? COUNTRY: United States of America
? ZIP: 94306-1840
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
? COMPUTER: PC
? OPERATING SYSTEM: DOS
? SOFTWARE: Wordperfect 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/415,751
? FILING DATE: 03-APR-1993
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/071,880
? FILING DATE: June 1, 1993
? APPLICATION NUMBER: 07/891,301
? FILING DATE: May 29, 1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Hana Dolezalova
? REGISTRATION NUMBER: 30,518
? REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 324-1677
? TELEFAX: (415) 324-1678
? INFORMATION FOR SEQ. ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 361 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? ORIGINAL SOURCE:
? ORGANISM: Cryptosporidium parvum
? FEATURE:
? NAME/KEY: Positions coded by nonsense codons are
? US-08-415-751-3

Query Match 4.7%; Score 112; DB 1; Length 361;
Best Local Similarity 38.3%; Pred. No. 0.0019;
Matches 23; Conservative 12; Mismatches 25; Indels 0; Gaps 0;

OY 382 NNCNTILNCNRSDDTTTNNNNNSVDRPSDSTNTNNNIYDHPNDINNKNNVYDNKDNNSD 441
Db 115 NNHNNSXSNHNNNSKNTNNNNNNKGSXKATNSHNNNNIKANSYNNNNKSNNNNNNSAND 174

RESULT 4
? Sequence 8, Application US/08914999
? Patent No. 6346406
? GENERAL INFORMATION:
? APPLICANT: Ryazanov, Alexey G.
? APPLICANT: Pavut, Karen S.
? APPLICANT: Hailt, William N.
? TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
? TITLE OF INVENTION: AND METHODS OF USE THEREFOR
? NUMBER OF SEQUENCES: 25
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: David A. Jackson, Esq.
? STREET: 411 Hackensack Ave, Continental Plaza, 4th
? STREET: Floor
? CITY: Hackensack
```

```

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostellium discoideum
US-08-914-999-8

```

Query Match 4.4%; Score 105; DB 4; Length 732;

Best Local Similarity 20.1%; Pred. No. 0.033; Matches 76; Conservative 61; Mismatches 167; Indels 74; Gaps 17;

```

OY 117 EFKLE-EXQIVNVSVKLNSFIIEEG---SDDKFEPPSLCSKPRKRORGRNNTRL 172
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 34 EKKIRFPYGTQSOTLQIGKMLPSGGGATADSKEFKKARNTLADIQYKVDLYVAV 93
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 173 KV-----CFLPLDSPSLTNGENGTITLNDGNRLGYPEATELAGOFEMTSNI----PP 222
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 94 KSKRTNDSLEPTLIALPLDGSERAIKWEYD-----PYTF--TAQWTCITLVKVEPV 144
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 223 AIAHSLDAGAKVILTSEAVVPATKRLSAERSEARSHLLQKROF-----YHSR 274
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 145 PPAEAFKRAYHTLDSKSGASGRVYSKIGKKPTRPSTFEDYKQMLAKKADYNSRK 204
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 275 VOPMALEOVMS-----DRDSEDEV--DDDVADFEEDROMLDFVYNKDEK---QFMHL 322
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 205 -PPKRIEFLQSCVLEFVDRSTSDLICGAEPYEGQYRKYNNSGFSVNDERNTPOSFHF 263
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 323 WNSFVRKORVLTAD-----GHISWACFAFSRFEYKEKELHRYSSLFCWCRFLIKIMNGIVDS 378
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 264 TYEHSNHOLLITDIDGVC-----DHYTDPQIHITYDGV---GFGIGNLGQKGEFEK 310
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 379 ATINNCNTILE--NCRNSDPTTTNNNSVDRPS-----DSNTNNNTIVDHPNDIN 427
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 311 LDTHCMAICQYLNIQSLINPKSEKSDCGTVPRPDLIFPTSERDNNNNNNNN--NNNNNNN 368
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 428 KKNVNDKDNNSRDKVIK 445
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 369 NNNNNNNNNNNNSISK 386
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

RESULT 5

US-08-700-152A-4
; Sequence 4, Application US/08/00152A
; Patent No. 5994622

GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamura, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 4

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,152A
FILING DATE: 20-AUG-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067200US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-700-152A-4

```

Query Match 4.3%; Score 103; DB 2; Length 432;

Best Local Similarity 21.5%; Pred. No. 0.023; Matches 88; Conservative 54; Mismatches 134; Indels 134; Gaps 20;

```

OY 74 DCNNTLOKTEVEREDSCPC-----SMLGSPFKGLQFHLNSSHDLFEFEFKLEEXQIVNV 129
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 3 DLNDAPHQHQ--REESEECYSSPSKRVGSFS-----NSS----- 36
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 130 SVKLNSFIIEEGSDDDKFEPPSLCSKPRKRORGRNNTRLKACFLPLDSPSLTNGTE 189
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 37 ---SSAVYIEDGSDDEL-----NRRAPNPVLTQHOFPEMDS----- 71
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 190 NGITLNDGNRLGYPEATELAGOFEMTSNIPPALHSLDAGAKVILTSEAVVPATKTR 249
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 72 -----NGCGVASGFPRAHMPGVKFCQSLATGSSAGKATNVAHV---EQAQPLKSR 122
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 250 KLSAERSEA-----RSHLLQKROFY-----HSHRVOPMALEOVMSDRDSED 291
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 123 RGRPRRSQYRGVTFEYRTGRWESHIMDCGQVYLGEDPTAHAAARAYDRAAIKFRGVEA 182
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 292 EYDDVDVADDED--ROMLDFVYNKDEQOFMLNMSFVRKORVITADGHISWACFAFSRY 349
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 183 DTFENIDVDLDDDKQM-----TNLTKEEFVHV---LKRQST-----GPPRGS 221
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 350 EKELHRYSSLFCWCR-----LFLIKIMNH--GLVDSATINNCNTILENCR-----N 393
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 222 SK--YRGVTLHKGCGWEARMGQFLGKYYLYGLFD-----TEVEARADKKAIKCN 271
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 394 SSDTTTNNNSVDRP--SDSNTNNNTIVDHPNDINKNVNDKDNNSRD 441
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 272 GQDAVTNEDPSIYDEELNASSGNGPTTPQDHMLDLSLGNSANSK--HKSQD 320
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

RESULT 6

US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480

GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay

Fri Jun 20 09:03:39 2003

us-09-890-220-2.ra1

Page 4

FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/541.782
CURRENT FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1073
TYPE: PRT
ORGANISM: S.pombe
US-09-541-782-6

Query Match
Best Local Similarity 18.3%; Score 103; DB 4; Length 1073;
Pred. No. 0.1;
Matches 108; Conservative 75; Mismatches 200; Indels 206; Gaps 24;

```
OY 13 EEVISTDE-----NLIYCKPVRLYNIFHLRSIGNPSFLPRCLNYKIGAKR 58
DB 394 EETISTLEVAARAKSIRKPNQNLVFRKVLKIDVLDIRLKN-----LN-----A 441
OY 59 KRKRSYGVVYVYKDCNNLTOKTEYREDSCPCSMGSGFKGLQFHLNSHDLFEFEF 118
DB 442 TRKKNVYLAESTYKELMDRYQNKDL-----LCQEQARKLEVLDLVANKSSRQLQYVS 494
OY 119 KLPEEY-----QVNVSVKLNSEIFEEGSDDD-----KEEP-----150
DB 495 KSNQEHKEVEALQDLVNSSTELSVKSENEKLNELVLEIKRKKEETNEAKITTYAT 554
OY 151 -----FSLCK--PKRRQGRGNTRRLKYCFPL--DSPSLNGTNGCI 192
DB 555 DLQYRESKEYIASLYEKIDRTERNNKNNENFNWLNKFNLLTMRSPFGSFTDETNGYF 614
OY 193 TLNDGNRGIG-----YPEATLAGQFE-----MTSNIPPAIAHSSLDGAKVI 236
DB 615 TLNDPNSMRELLNTHSNOLLISMTKITEHFQSLDEALQARSSCAVNSSL-----L 669
OY 237 LTSE-----AVPA-----TKTKLSAERSEARSHLLQKROFY 270
DB 670 IVELDKSKNSLDLALHSLDIODISMSSQKLGNGISSELELODKMESYRQLVQELRSY 729
OY 271 ---HSH-RVQPMALQVMSDRDS-----EDEVDVADFEEDROMLDFV 310
DB 730 NLQHTHEESQKELMYGVANDIDALVKTCTTSLNDADILSDYISDQSKFESKOO-DLIA 788
OY 311 DVNDEKQFM-----HL--WNSFVKORVIADG-----336
DB 789 NIGKIYVSNFLOEONESLYTKADILHSHLNDTNSNIRKANEIMNNNSEEFELRNASQAIEV 848
OY 337 ---HSMACEAFSRFE--KELHRYSLFCWRLFLIKLMNHGLVDSATINNCTIL 388
DB 849 GANKERIOKTVENGSQLDSSKSKAIHNSRSRMYDCHLALAESQKGV--NLEVOQLDRL 906
OY 389 ENCRRNSDPTT-----NNNSVDRPSDSNTNNNNIYDH 422
DB 907 OKVKEHSDNTKEKHQQLDLLESLVGNNDNLDSIKTPHTELQKTIKH 955
```

RESULT 7
US-09-723-820-6
Sequence 6, Application US/09723820
Patent No. 6468760
GENERAL INFORMATION:
APPLICANT: Nislow, Corey
APPLICANT: Berard, Christophe
TITLE OF INVENTION: Antitungal Assay
FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/723,820
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 09/541,782
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6

LENGTH: 1073
TYPE: PRT
ORGANISM: S.pombe
US-09-723-820-6

Query Match
Best Local Similarity 18.3%; Score 98.5; DB 4; Length 1388;
Pred. No. 0.1;
Matches 108; Conservative 75; Mismatches 200; Indels 206; Gaps 24;

```
OY 13 EEVISTDE-----NLIYCKPVRLYNIFHLRSIGNPSFLPRCLNYKIGAKR 58
DB 394 EETISTLEVAARAKSIRKPNQNLVFRKVLKIDVLDIRLKN-----LN-----A 441
OY 59 KRKRSYGVVYVYKDCNNLTOKTEYREDSCPCSMGSGFKGLQFHLNSHDLFEFEF 118
DB 442 TRKKNVYLAESTYKELMDRYQNKDL-----LCQEQARKLEVLDLVANKSSRQLQYVS 494
OY 119 KLPEEY-----QVNVSVKLNSEIFEEGSDDD-----KEEP-----150
DB 495 KSNQEHKEVEALQDLVNSSTELSVKSENEKLNELVLEIKRKKEETNEAKITTYAT 554
OY 151 -----FSLCK--PKRRQGRGNTRRLKYCFPL--DSPSLNGTNGCI 192
DB 555 DLQYRESKEYIASLYEKIDRTERNNKNNENFNWLNKFNLLTMRSPFGSFTDETNGYF 614
OY 193 TLNDGNRGIG-----YPEATLAGQFE-----MTSNIPPAIAHSSLDGAKVI 236
DB 615 TLNDPNSMRELLNTHSNOLLISMTKITEHFQSLDEALQARSSCAVNSSL-----L 669
OY 237 LTSE-----AVPA-----TKTKLSAERSEARSHLLQKROFY 270
DB 670 IVELDKSKNSLDLALHSLDIODISMSSQKLGNGISSELELODKMESYRQLVQELRSY 729
OY 271 ---HSH-RVQPMALQVMSDRDS-----EDEVDVADFEEDROMLDFV 310
DB 730 NLQHTHEESQKELMYGVANDIDALVKTCTTSLNDADILSDYISDQSKFESKOO-DLIA 788
OY 311 DVNDEKQFM-----HL--WNSFVKORVIADG-----336
DB 789 NIGKIYVSNFLOEONESLYTKADILHSHLNDTNSNIRKANEIMNNNSEEFELRNASQAIEV 848
OY 337 ---HSMACEAFSRFE--KELHRYSLFCWRLFLIKLMNHGLVDSATINNCTIL 388
DB 849 GANKERIOKTVENGSQLDSSKSKAIHNSRSRMYDCHLALAESQKGV--NLEVOQLDRL 906
OY 389 ENCRRNSDPTT-----NNNSVDRPSDSNTNNNNIYDH 422
DB 907 OKVKEHSDNTKEKHQQLDLLESLVGNNDNLDSIKTPHTELQKTIKH 955
```

RESULT 8
US-09-572-191-2
Sequence 2, Application US/09572191
Patent No. 6355466
GENERAL INFORMATION:
APPLICANT: Berard, Christophe
APPLICANT: Sakowicz, Roman
TITLE OF INVENTION: No. 6355466 motor proteins and methods for
TITLE OF INVENTION: Their use
FILE REFERENCE: 1017
CURRENT APPLICATION NUMBER: US/09/572,191
CURRENT FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1388
TYPE: PRT
ORGANISM: Human
US-09-572-191-2

Query Match
Best Local Similarity 19.0%; Score 98.5; DB 4; Length 1388;
Pred. No. 0.46;

Fri Jun 20 09:03:39 2003

us-09-890-220-2.rai

Page 6

```
; Patent No. 6310046
; GENERAL INFORMATION:
; APPLICANT: Patrick E. Duffy
; APPLICANT: Christian F. Ockenhouse
; TITLE OF INVENTION: SEQUESTRAIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John Moran
; STREET: USA MRMC - MCMR-JA
; CITY: FORT DETRICK, FREDERICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,896B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, John
; REGISTRATION NUMBER: 26,313
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-559-896B-2

Query Match          4.1% Score 98; DB 4; Length 652;
Best Local Similarity 21.8%; Pred No. 0.15;
Matches 41; Conservative 34; Mismatches 71; Indels 42; Gaps 7;

QY 265 OKROFYSHRQPMALQVMSDRSDEY-----DDVADPEDROMLDFVYVNDK 315
DB 289 EEKEKTHREKLEKIEKINKMDQIDKITEELINKMSDEICQVRRALIE--DIQKE 345
QY 316 EKQPMILMSFYRKQRYADGHISWACEAFSRYEKELEHRYSSLFQCMRLFLIKLNNHGL 375
DB 346 KIQNLEL-----EIDRLTYEELDRMR---EAREIIPMRN-- 378
QY 376 VDSATINNCNTILENCRSSDTTNNNSVDRPSDSTNNNTVDHPNDINKNNVNDK 435
DB 379 --LSRNEKNDITHRIKINKESNOKNKEVNVFIIDNNDNSNNNNNNRRVNNLNN--KHT 435
QY 436 DNNSRDKV 443
DB 436 NNNYENNV 443

RESULT 12
US-08-446-855A-2
; Sequence 2, Application US/08446855A
; Patent No. 5849573
; GENERAL INFORMATION:
; APPLICANT: Stewart, Thomas S
; APPLICANT: Flores, Maria Y
; APPLICANT: O'Sullivan, William J
; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
```

```
; STREET: 1100 No. 5849573th Gleebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,855A
; FILING DATE: 06-Jul-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 47-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4100
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2391 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-446-855A-2

Query Match          4.0% Score 94; DB 2; Length 2391;
Best Local Similarity 21.4%; Pred. No. 3.4;
Matches 44; Conservative 40; Mismatches 64; Indels 58; Gaps 12;

QY 266 KROFYSHRQPMALQVMSDRSDEY-----FVDDVADPEDROMLDFVYVNDKQFQFHL 322
DB 269 KEFNYTN-----EMITDSKSEHDNDINSISFNKCSISSEF-DKSKRYVINH- 319
QY 323 WNSFYR-KQRYADGHISWACEAFSRYEKELEH-----RYSLEFQCMRLFLIKLNN 372
DB 320 --TLRDKMLNITS-----SEETLDLHNCFSNDSQDSFF--KLIGICEDYD 364
QY 373 HGLVD-----SATINNCN-----TILENCRSSDTTNNNSVDRPSDSTNNNT--- 419
DB 365 KYLDLEBNASFHYNNVDEYGYDVNNKNTNIIISNKKIEQNNNNNNNNNNNEVDYI 424
QY 420 -VDHPNDINK-----NNVDNKN 437
DB 425 KKDEDNNVNSKVFSQYNNNAONNEH 450

RESULT 13
US-09-150-741-2
; Sequence 2, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PI6380
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: A093/00617
; EARLIER FILING DATE: 1995-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2391
```

```

? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: protein
US-09-150-741-2

```

Query Match	4.08;	Score 94;	DB 4;	Length 2391;
Best Local Similarity	21.48;	Pred. NO. 3.4;		
Matches 44;	Conservative 40;	Mismatches 64;	Indels 58;	Gaps 12;

QY 266 KRQFYHSRHYQPMALQVNSDRDSED---EYDDVADFEEDRQMLDFVDVYDKDEKQPMHL 322Z

D6 269 KEENNYTN-----EMITNDSMEDHDNEINGSISFNKPSISGF-DKESKKNVYNH- 319

```

QY      323 WNSFVR-QQVINDGHISNACEAFSEFYEEHL-----RYSLEFWCRRLFLILMN 372
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      320 --TLRLDRKMLITS-----SEELYKDLHMCNFSNSSDKNDSEFF--KLYGICEYD 364

```

```

Oy 3/3 HGVD-----SATINCN-----TLENCINSSDTTTTNNNNSVDPDSNTNNNNI--- 419
      | : | | : : | | : : | | : : | | :
Db 365 KYLIDLEENASFHINNVDGYDYVKNKTILSNKKIEQNNNNNNNNKNNKNNNNNEYDYI 424

```

```

qy      420 -VDHPNDINNK-----NNVDKNKN 437
          | |:::| | |::
Db      425 KKEEDNNVNSKVFYSQYNNNAQNNHH 450

```

RESULT 14
US-09-384-162-6
: Sequence 6. Application US/09384162

```

: Patent NO. 03/0/04/
:
: GENERAL INFORMATION:
:
: APPLICANT: Xing, Ti
:
: APPLICANT: Malik, Kamal

```

: APPLICANT: Miki L., Brian
 : TITLE OF INVENTION: No. 6376747el Plant-Derived Map Kinase Kinase
 : FILE REFERENCE: 08-884280US

```

; SOFTWARE: Patentin Ver. 2.0
; NUMBER OF SEQ ID NOS: 24
; CURRENT FILING DATE: 1999-08-27
; SEQUENCE INFORMATION: REF ID: A01106

```

```

: LENGTH: 286
: TYPE: PRT
: ORGANISM: Dictyostelium discoideum

```

Query Match 3.9%; Score 92; DB 4; Length 286;
Best Local Similarity 37.0%; Pred. No. 0.18;

[illegible]

RESULT 15
US-09-177-240-2

```

; Sequence 2, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert T.

```

APPLICANT: Umad, Nit
APPLICANT: Kiyosue, Tomohiro
APPLICANT: Yadehari, Ramin
APPLICANT: Marcossian, Linda

: APPLICANT: malaga, john
 : APPLICANT: Goldberg, Robert B.
 : APPLICANT: The Regents of the University of California
 : TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit

FILE REFERENCE: 023070-086120US

```

: CURRENT APPLICATION NUMBER: US/09/177,249
: CURRENT FILING DATE: 1998-10-22
: EARLIER APPLICATION NUMBER: US 09/071,838
: EARLIER FILING DATE: 1998-05-01
: NUMBER OF SEQ ID NOS: 324

```

```

; SEQ ID NO 2
;
; LENGTH: 689
;
; TYPE: PRT

```

US-09-177-249-2	
Query Match	3.88; Score 91.5; DB 4; Length 689;

Matches 49; Conservative 39; Mismatches 61; Indels 93; Gaps 11;
246 TTTTTLKLSERSEARSHLLQKRFYHSHRYQPMALQVMSDRSEDEYDDVDVADFEDROM 305

Dd	142	TKSSQLMAE-----SDSVIGKRLIYY-----LNGEAL-LSSEDEDEDEDE-----	163
Qy	306	LDDFVDVAKDEKQFMHLMNSFVKQRVITADGHISWAC-----EAFSRFYKE-	352

D6 184-----EEIKKEKCEFESEDVRF-----WTVGQDYGIDDLVVRALAKYLEVDV 227

QY 353 ---LHRYSLFWCWRLEFLIKLMNHGJVDSATINNCCNTIL-----ENCRN----- 393

DB	228	SDILERYNEL-----	KLKNDGTAGC	EASDLSK	ITTTAF	QFQFADRR	CHRCRM	IFDCH	278
QY	394	-----SSDTTNNNSVD	RPSD-----	-----	-----	-----	-----	-----	435

DB 279MMHEKYEPESRSSSEDKSSLEFEDENQPCSEHCYKLVRSYTEADHVMDNPNNSISNKRIVSDP 338

OY 436 DN 437 : 1

Search completed: June 13, 2003, 15:45:55

FOR LINES 1-40000

Search completed: June 13, 2003, 15:45:55
Job time : 28 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:34:30 ; Search time 25 Seconds
(without alignments)
1903.099 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRQNCRAKSSPEVISHDE.....INKNKNVDKNKNSDKYIK 445

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCRUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	9.5	388	10	US-09-764-864-931
2	221	9.3	739	9	US-09-874-162A-5
3	221	9.3	776	9	US-09-874-162A-8
4	153	6.4	289	10	US-09-764-864-911
5	127.5	5.4	961	10	US-09-801-368-132
6	126.5	5.3	2150	9	US-10-135-322-17
7	125.5	5.3	1501	10	US-09-924-154-17
8	124	5.2	1288	9	US-10-087-780-2
9	122.5	5.2	1331	9	US-10-087-464-50
10	115	4.8	175	10	US-09-764-864-1355
11	113	4.8	807	9	US-09-820-843A-108
12	111.5	4.7	1143	10	US-09-924-154-14
13	105	4.4	732	10	US-09-832-292-12
14	105	4.4	732	10	US-09-994-485-8
15	104.5	4.4	97	9	US-09-858-935B-77
16	104.5	4.4	666	10	US-09-801-368-36
17	101	4.2	534	10	US-09-801-368-124
18	100.5	4.2	97	9	US-09-858-935B-72
19	99.5	4.2	292	10	US-09-764-864-1369

20	99.5	4.2	861	9	US-09-820-843A-109	Sequence 109, App
21	98	4.1	97	9	US-09-858-935B-76	Sequence 76, Appl
22	98	4.1	652	10	US-09-351-794A-2	Sequence 2, Appl
23	97.5	4.1	97	9	US-09-858-935B-73	Sequence 73, Appl
24	97.5	4.1	97	9	US-09-858-935B-74	Sequence 74, Appl
25	97.5	4.1	97	9	US-09-858-935B-75	Sequence 75, Appl
26	97	4.1	440	9	US-10-083-357-1266	Sequence 1266, Ap
27	97	4.1	559	10	US-09-801-368-256	Sequence 256, App
28	96.5	4.1	97	9	US-09-858-935B-70	Sequence 70, Appl
29	96.5	4.1	97	9	US-09-858-935B-71	Sequence 71, Appl
30	95	4.0	1093	10	US-09-801-368-392	Sequence 392, App
31	94.5	4.0	393	9	US-10-153-668-76	Sequence 76, Appl
32	94.5	4.0	339	9	US-10-153-668-78	Sequence 78, Appl
33	94.5	4.0	556	9	US-10-012-896-1005	Sequence 1005, Ap
34	94.5	4.0	1610	9	US-10-155-533-9	Sequence 9, Appl
35	93.5	3.9	497	9	US-09-820-843A-32	Sequence 32, Appl
36	93	3.9	676	10	US-09-801-368-302	Sequence 302, App
37	93	3.9	2086	10	US-09-815-242-5639	Sequence 5639, Ap
38	93	3.9	5795	10	US-09-815-242-12610	Sequence 12610, A
39	92.5	3.9	704	10	US-09-801-368-218	Sequence 218, App
40	92.5	3.9	758	10	US-09-801-368-224	Sequence 224, App
41	92.5	3.9	762	9	US-09-738-628-4825	Sequence 4825, Ap
42	91.5	3.8	91	9	US-09-858-935B-69	Sequence 69, Appl
43	91.5	3.8	689	10	US-09-071-838-2	Sequence 2, Appl
44	91.5	3.8	1786	9	US-09-742-096-3	Sequence 3, Appl
45	91	3.8	1051	10	US-09-911-888-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-764-864-931
Sequence 931, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 931
LENGTH: 388
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-864-931

Query Match	9.5%	Score 225	DB 10	Length 388
Best Local Similarity	21.1%	Pred. No. 6, 2e-13		
Matches	84	Conservative	64	Mismatches 151; Indels 100; Gaps 12
58	RRRK-----SRSTGVVFNFKDNNLTQLEFREDCSCFCSMLGSGFKGLQPHLNSHD	112		
62	REKQTPNNRNKRLNLFYGLNNNTROTARDDLHCWCTLNCRLKLSLKLKILCHS	121		
113	LEFEFKLEFETQYVNVSVKLSLFEEDSGDDKFEPSLCSKPRK-RRQNG--GRNN	168		
122	RIFFVYVHPKARDVSI-----NKGYG-----SYAGNPQDIHROGFAFSRNG	167		
169	-TRRLKVCPLRPSLSTENGTLINDGSRGLGYPRATLAQFEKTSNIPALNHS	227		
168	PKRPIPIHILICRKRRTKASSEPLESDG-----	198		
228	SIDAGAKVLTSEAVPATKTRKLSERSEARSHLLQKROPYHSHRVOPMALQVMSDR	287		
199	-----EVEQQRYSYSGH-----NRLTFHNDTGLPLRQDM--EV	230		
288	DSEDEVDDVADPEDROMIDFVDVNVKDEKQPMHLNLSFVRKQRYIADGHSWACEAFSR	347		
231	DSEKDEDEMLEKRTITTOIEEFSDVNKEGEYMKLMLNLHVMKHGFIADNOMHACMLEVE	290		

Fri Jun 20 09:03:39 2003

us-09-890-220-2.rapb

Page 2

```
QY      348 FYEKEIHRYSLEWCWCRILYLKLMHNGLVDSATINCNTIEENCRNSDPTTTNNNNSVD    407  
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::  
Db     .   291 NYGQKIIR-KNLGRNMMLVSGMHPNPULSISIMSDIKATKLREMG-----OKLE    338  
  
QY      408 RPDSMTNNNNIVDPHN-----DIINKKNAYDNKDNNSS    439  
          :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db       339 KGESAPANEETFEONGTANGFSELSNEKKALETDVS    377
```

```

RESULT 2
US-09-874-162A-5
Sequence 5, Application US/09874162A
Patent No. US20020155452A1
GENERAL INFORMATION:
APPLICANT: Koonetz, Jason
APPLICANT: Sklar, Jeffrey
TITLE OF INVENTION: FUSION OF JAGG1 AND JTAZ1 GENES IN
TITLE OF INVENTION: ENDOOMETRIAL STROMAL TUMORS
FILE REFERENCE: 05311-024001
CURRENT APPLICATION NUMBER: US/09/874,162A
CURRENT FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: US 60/209,093
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 739
TYPE: PRT
ORGANISM: Homo sapiens
US-09-874-162A-5

```

Query Match	Score	DB	Length
9.38;	221;	9;	739

Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

```

OY 58 RKRK-----SRSGMVFYVYKDCNNNTLQATEVREDCSCSPFCMSLGSCKGLOPLNSHD 112
Db 413 RKREKDPNEROKRLRIFYOGLYNNNTROOTERDRDLHCPCWCTLRCRLYSLLKHLKICHS 472
OY 113 LFEFEFLFREYOTVAVVXLSNFIJEEGSDODKFEDEPISLCSMPK-RRORG---GRNN 168
Db 473 RFJFNYYVHKGARIDVSI-----NCCYDGC-----SYAGNPDIOHRRPGPFRFRNG 518
OY 169 -TRRLKCFPLDPSLJTGTEGTITLNDGRGLGYPEATLGAOREMNSIDPALHS 227
Db 519 PVKRPITLHLCVRPKTKASMFLESDG-----549
OY 228 SLDAGAKVILTSAAVVPARTKRLKLSAERSEASHLLLRQFYSHHYORPALEQVYMSDR 287
Db 550 -----EVEDQKRTSSGK-----NRLFESDTCPLPRPDEM--EV 561
OY 288 DSEDEVDVDYADEEDRQMLDDFVYVYKDEKQPMHILMNSFVKORYIADGHSIWAACEFSR 347
Db 582 DSEDEKPEPELREKRTTQQLTEEFSDVNBEGEKVEMKLMNLYHVKHGFIDADNOMNHCMJLFE 641
OY 348 FYEKELHRSSLEFCWRPLFLKLMNHGLVDSATINCCNTILMCRNSSDPTTTNNNSVD 407
Db 642 NYQOKITIK-RNLCRRNMLHLYSMHPFNILSTISDIAVTKLRQMO-----OKLE 689
OY 408 RPSDSNTNNNIVDHN-----DINKKNVYDKRNS 439
Db 690 KGESAPANEETIEONGTANGFSSINSKERALETDSVS 728

```

RESULT 3
US-09-874-162A-8
; Sequence 8, Application US/09874162A
; Patent No. US20020155452A1

```

; APPLICANT: Koontz, Jason
; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAZF1 AND JAZ1 GENES IN

```

```

1  TITLE OF INVENTION:  ENDOMETRIAL STROMAL TUMORS
2  FILE REFERENCE:  05311-024001
3  CURRENT APPLICATION NUMBER:  US/09/874.162A
4  CURRENT FILING DATE:  2001-06-04
5  PRIOR APPLICATION NUMBER:  US 60/209,093
6  PRIOR FILING DATE:  2000-06-02
7  NUMBER OF SEQ ID NOS:  23
8  SOFTWARE:  FastSeq for Windows Version 4.0
9  SEQ ID NO 8
10 LENGTH:  776
11 TYPE:  PRT
12 ORGANISM:  Homo sapiens
13 US-09-874-162A-8

```

Query Match	9.38;	Score 221;	DB 9;	Length 776;
Best Local Similarity	31.18;	Prod No 4	10-13.	

Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

Qy	58	RRKK-----SRGTGVVFNVDKNNITLOKTIVREDCSCPFMSICSPFGJLDFHLEMSSD	112
Db	450	RKKEDPNNENRÖKRLRIFÖFPLYNNNTROQTARDDJLCPCMTJCNCKRLKLSLKHLKICS	509
Qy	113	LFEFEERKLFEEÖTAVNSVKNLSNFFEEEGSDDDKFFEPFLSCSPKK-RRÖRG-----GRNN	168
Db	510	RFFENYVYHPKRGARHDVSI-----NECYDG-----SYAGNPQDHRÖGFAPFRNG	555
Qy	169	-TRRLKVCGLPLDPSLNGTENGJTTLLNDNRGLGYREATELAQFEMTSMIPALHNS	227
Db	556	PVKRPITTHLVCRRKRTKASMSERLESDS-----	586
Qy	228	SUDAGKAVLTSEAVPAPTRKRLSAPSEARSLLÖKROFVHSHRPOVMALÖEYMSDR	287
Db	587	-----EYDÖQRITYSGGH-----NLYHSHPTCLPARKÖEM--EY	618
Qy	288	DSEDEYDDVDALFEDERÖMLDPOVYNNKDEÖFPHILMNSPVKROKRYADGHISMACEPSR	347
Db	619	DSDEDEPEMLREKTIITÖIEERSDÖNBEKRYKMLNLYHMKGFTADÖNNHACHLFE	678
Qy	348	FYKREKLHARYSLFWCWRFLPIKLNHGLVDSATINNCNTYLENCRRNSDPTTTNNNSVD	407
Db	679	NTGÖKILIK-KALCNRFMLHLVSMHDPFLMISIMSDAXVTLKREMO-----ÖKLE	726
Qy	408	RPSDSNTNNNITVDHPN-----DINKNNVNDKONNS	439
Db	727	KCESASPAHEETIEÖNGTANGFELSINKERALETÖDVS	765

1
E
C
T
I
E

S-09-764-864-911

Patent No. US20020132753A1

APPLICANT: Rosen et al.

FILE REFERENCE: PT223

CURRENT FILING DATE: 2001-01-17

NUMBER OF SEQ ID NOS: 1792

SEQ ID NO 911

TYPE: PRT

ORIGINATOR: HOME DEPOT
S-09-764-864-911

Query Match	6.48; Score 153; DB 10; Length 2
-------------	----------------------------------

Best local similarity 25.5%, Freq. no. 2.7e 00,
Matches 49; Conservative 43; Mismatches 89; Indels

245 ATKTRKLSAERSEA--RSHLLQKRQFYHSHRVQPMALQVMSDRDS

```
88 ASMSEFLESEDGEVEEOORTYSSGHNRLYFHSDTCLPLRPOEM--EVDSD
```

```
QY 302 DROMLDFVYNKDEKOFHMLNMFVRKQKVADGHSWACEAFSRFYEKELHRYSLEFM 361
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 146 TITQIEFSDVNEGEVEVKMLNLMVAKHGFIADNOMACMLFVENVGKLIK-KNLOR 204
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 362 CWRFLFLIKMLNGLVDSATINNCNTLLENCRNSSDTTNNNSVDRPSDSNTNNIYD 421
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 NFWLHLVSMHDFLJISMSIDKAVTKLREMQ-----QKLEKGSASAPANEITE 253
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 422 HRP-----DINNKNNVNDKNNS 439
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 EQNGTANGFSEINSKEKALETDSVS 278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-09-801-368-132
; Sequence 132, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-132

Query Match
Best Local Similarity 41.3%; Pred. No. 0.0041;
Matches 26; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

QY 382 NNCNTLENCRNSSDTTNNNSVDRPSDSNTNNIYDHPNDINN-KNNVDKNNSR 440
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 501 NNNNNNDNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 560
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 441 DKY 443
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 561 DNT 563
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
US-10-135-322-17
; Sequence 17, Application US/10135322
; Patent No. US2002017301A1
; GENERAL INFORMATION:
; APPLICANT: BENFEY, PN
; APPLICANT: HELARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AMM
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
```

```
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17

Query Match
Best Local Similarity 17.5%; Pred. No. 0.017; Length 2150;
Matches 85; Conservative 64; Mismatches 175; Indels 163; Gaps 16;

QY 2 CROCRKAKS-----SPEEVISTDENLITCYKPYALYNIFHRSIGN-----PSFLPCLNY 52
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 353 CGSNGSGNGIGPLSPRLNLSLNGVAVSPRNI-----HLNNLNNSNLPPLSPRHINF 406
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 53 KIGAKRRKRSRTGAVYVYKDCNNTLQTEVREDCSCFPCSMLCGFKGLQPHLSSHD 112
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 407 HINVSMLNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 113 LFEFEFKLFEEVQTVAVSKLNSFIEEBSGDDDKFEPPLSKPKRRROGR----- 166
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 436 -----NVSRRNNHNIISPRGSN---ISP-----RSNNGGSTITISR 468
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 -----NNTRLKVCFLPLDSPSL--TNGTENGITLLDNGRGLGYPEA-TELAGO 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 469 NISNNNNIINNINNINILPRLNPSRLEVNPTNSRLATLSNLTPLVSSLTSSNNN 528
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 EMTSNIPALAHSS-----LDGAKVILITSEAVV--PATKTRKLSAESEARSHLLQK 266
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 529 QSNNTNTPSINNNGNGHGIOTISEILGNKVVYVNNNNNNNNNNNNNNNNNN 588
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 267 RQYSHRYQPMALQVMSDROSEEDVDDVADFEEDRML--DDFYD-----VYK 314
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 589 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 648
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 315 DEKOFHMLNMFVRKQKVADGHSWACEAFSRFYEKELHRYSLEFMKMLNNG 374
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 649 DDND-----ENDG 656
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 375 LVDSATINNCNTLLENCRNSSDTTNNNSVDRPSDSNTNNIYDHPNDINNKN 434
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 657 NSNNTSNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 716
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 435 KDNNSRD 441
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 717 NNNNNNN 723
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-09-924-154-17
; Sequence 17, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Natum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1501
; TYPE: PRT
; ORGANISM: Mammalian
```

```

US-09-924-154-17
Query Match 5.38; Score 125.5; DB 10; Length 1501;
Best Local Similarity 19.28; Pred. No. 0.012; Indels 127; Gaps 13;
Matches 83; Conservative 59; Mismatches 164; Indels 127; Gaps 13;

QY 48 RCLMYKIGAKRRKRSR-----TCMVYNYVOCNNYLOKTEV-----RECS 89
DB 670 RKLAYELOSRYKDKRDLKFLSLADKNTTFLKNNANCSN-IDFTIFOLDLKEKCS 728
QY 90 CPFSMIGSGFSGOGLFNLSSHDLEFEERKLEFYTYTN-VSYKLSNLFEEEGSDDKF 148
DB 729 CMTQVLELVNKEMLSDNSBDATDISEKNGEELVYHNNSVYASGKKELEKSKDE- 787
QY 149 EPFSLCKPRKRRRGRNNTRLKYCFLPLDSPSLTNGEITLLDNGNGLGIPAT 208
DB 788 -----QPEKAKQ-----TNGITL-VRTDKDSNNKGRDYAT 818
QY 209 ELAQFEMTSNIPALAHSLDAGAKYILTSEAVVATRTKLSAERSARSHLLQKQ 268
DB 819 D-----TKNSPENLKYQEHGNGETIKEEPKLPESSETLOSQDLAEKQKQKEE 871
QY 269 FYSHRYOPALBEVMSDRSE--DEVDVADDFEDRQMLDFVYNKDEKQFHLNMF 326
DB 872 PKKQEEPKKQEEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 927
QY 327 VRKQVYADGHISMACEAFSRFYKEKELHRSLSFWCWRFLIKLNHGLVDSATNNCT 386
DB 928 ASEQNEISSGQ-----PNDINKNNV 432
QY 387 ILENCRRNS-----DTTNNNSVDPRSPDSNTNNNIVDH-----PNDINKNNV 432
DB 939 -EQVWSSPEVYPOETTSNGSSQDTKISSSTEPENNSVDATSMNLDPEKVENMS 997
QY 433 DNKNNSRDYIK 445
DB 998 DPTNTEPDASLK 1010

RESULT 8
US-10-006-780-2
; Sequence 2, Application US/10006780
; Publication No. US20030104496A1
; GENERAL INFORMATION:
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; APPLICANT: Guo, Jun
; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
; FILE REFERENCE: CYTOPO83
; CURRENT APPLICATION NUMBER: US/10/006,780
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1288
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-006-780-2

Query Match 5.2%; Score 124; DB 9; Length 1288;
Best Local Similarity 21.4%; Pred. No. 0.014;
Matches 78; Conservative 57; Mismatches 129; Indels 100; Gaps 17;

QY 110 SHDLEFEERKLEFYTYTNVYKLSNFI-----EEGSDDKFEPFSLCKPRKRRRG 164
DB 215 SHALNIDLK-----DINKNSLGIATIDLAGSRGAD-----TVSONKQOTFD 259
QY 165 GRNNTF--RLKYCFLPLDSPSLTNGEITLLDNGNGLGIPATEL-----AGQFE- 215
DB 260 GANINRSLALKECIRAMDS-----DKNHIFRDS-----ELFVLDIDIFVGSKS 305

US-09-764-1355
; Sequence 1355, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1355
; LENGTH: 175

US-10-087-464-50
; Sequence 50, Application US/10087464
; Publication No. US20030059436A1
; GENERAL INFORMATION:
; APPLICANT: Chisholm, Athar
; APPLICANT: Liu, David
; APPLICANT: Goel, Vikas
; APPLICANT: Goel, Vikas
; APPLICANT: Li, Xuepeng
; TITLE OF INVENTION: Band 3 Antigenic peptides, Malaria Polypeptides and Uses There
; FILE REFERENCE: S1237/7019
; CURRENT APPLICATION NUMBER: US/10/087,464
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Version 3.0
; SEQ ID NO 50
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-464-50

Query Match 5.2%; Score 122.5; DB 9; Length 1331;
Best Local Similarity 39.2%; Pred. No. 0.02;
Matches 29; Conservative 12; Mismatches 30; Indels 3; Gaps 2;

QY 371 WNHGLVDSATINN--CNTILENCRRNSDTTNNNSVDPRSPDSNTNNNIVDHPN-DIN 427
DB 595 YNNGYNDSTDNGNGYNSNSYNNNEDNNNNNNNDENNNDENNNDNNNNYNNYGN 654
QY 428 NKNVNDKDNNSRD 441
DB 655 NNNNNNNKDNND 668

RESULT 9
US-09-764-864-1355
; Sequence 1355, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1355
; LENGTH: 175

```



```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (166)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1355
```

Query Match
Best Local Similarity 20.4%; Score 115; DB 10; Length 175;
Matches 46; Conservative 30; Mismatches 74; Indels 76; Gaps 8;

```

QY 76 NNTLQKTEYREDSCSPFCSMICGSPKGLQPHLNSHDLFEFEFKLFEYQIVNVSYKLNS 135
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 8 NNTROGTEARDLHCPWCTLNCRLKLYSLKLKLCBSRTFFVYVHPKCARIDVSI----- 63
QY 136 FTFEEBGSDDKFEFPISLCSKPRK-RRQRC--GRNN-TRRLKVCFLPLDPSLNGTEN 190
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 64 ---NECYDG-----STAGNPQDIHROPFAFSRNGPYKRPITHILVCRRPKTKASMS 113
QY 191 GITLNDGNGRLGYPEATELAGOFEMTSNIPALAHSSLDAGAKVILTSEAVVPATKTKR 250
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 114 ELESEDEG-----E 122
QY 251 ISAESEARSHLLQKROFYSHRYQPMALQVMSDRSEDEVDD 296
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 123 VEQQRYSYSGH---NRXYFHSDFCLPLRQEM--EVDSEDEKDEPE 162
```

RESULT 11

```

US-09-820-843A-108
; Sequence 108, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE REFERENCE: 063915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: hypothetical protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: g113845292
US-09-820-843A-108
```

Query Match
Best Local Similarity 4.8%; Score 113; DB 9; Length 807;
Matches 24; Conservative 14; Mismatches 24; Indels 6; Gaps 1;

```

QY 372 NNGIVDSATINNCNTILENCRSSDPTTNNNNNSVDRPDSNTNNNIYDHPDINKNN 431
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 157 NYGLKKKITLRLDKIGYNNENITLNNKNNLNKNNNYVDNRN-----NNNNKNN 210
QY 432 VDNKDNN 439
   ::|||::
DB 211 INNNNNN 218
```

RESULT 12

```

US-09-924-154-14
; Sequence 14, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Natrum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-14
```

Query Match
Best Local Similarity 4.7%; Score 111.5; DB 10; Length 1143;
Matches 94; Conservative 70; Mismatches 171; Indels 171; Gaps 24;

```

QY 2 CRONCR-----AKSPSEYI---STDENLLIYCKPVRLYNIPIHLSIGNPSFLPCL 50
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 664 CKNKSEYKMWIDLKKSEYKQVDKTKDNKKMY-----DNIDEVK----- 705
QY 51 NTKIGAKRRKRSRGVYFNKDKCNTLQKTEYREDSCSPFCMKGSPKGLQPHLN-- 108
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 706 NKEANVYLLKREKCKADVNEDEKIFNES--PNEYEDMK-----KCEBIK---YLNEI 753
QY 109 ---SHDLFEFEFKLFEYQIVNVSVKNSFEEBGSDD-----KEEPE 151
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 754 KYPKTKHDIYDD--TFSTPFGDGPISINANINEDQSGKDSNTGNSPVSHEPE 811
QY 152 S-----LCSKPRKRQRGGRNNTRLKVCFLPLDPSLNGTENGITLNDGNGRLG 203
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 812 SDAIIVERLSDGDESSERTG-----ILDINDSPVNNVNE----- 847
QY 204 YPATELAGOFEMTSNIPALAHSSLD--AGAKVILTSEAVVPATKTKLSERBARSH 261
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 848 VHDASNTQGSVNTSDITNGHSSSLNRTTNQDIDIKIGRSGNEQSDNQSSSHSDNSGS 907
QY 262 LLL-----QKROFYHS--HRYOPMALQVMSD-----RSEDEVDDVD 299
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 908 LITIGVPSBDNTQNTYDSQPHRDPNPALASLPDDKINIEIGFDSRSSENGRGDTTSN 967
QY 300 FEDRQMLDFVDVNVKDEKQFMHLMNSFVRKORYIADGHISWACEAFSFEYERELHRYSSL 359
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 968 THDVR-----TNVTSRRVNSH--DFINGMANNAH-----HOY----- 1001
QY 360 FWCWRLFILKLMNHGLVDSATINNCNTILENCRSSDPTTNNNNNSVDRPDSNTNNNI 419
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 1002 -----IYQIENNGIIT-----RCQESAGNSVYKDKMPKSNFSSEN- 1037
QY 420 VDHNDINKNNKVVNDKNNSRDKYK 445
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB 1038 -DHKNIOEYNSRDTK--RVREIIR 1060
```

RESULT 13

```

US-09-832-292-12
; Sequence 12, Application US/09832292
; Patent No. US20020177205A1
; GENERAL INFORMATION:
; APPLICANT: Ryzanov, Alexey
; TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND
; FILE REFERENCE: 601-1-098CIP
; CURRENT APPLICATION NUMBER: US/09/832,292
; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 09/632,131
```

Fri Jun 20 09:03:39 2003

us-09-890-220-2.rapb

Page 6

PRIOR FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 732
TYPE: PRT
ORGANISM: Dictyostellium discoideum
US-09-832-292-12

Query Match
Best Local Similarity 20.1%, Pred. No. 0.38;
Matches 76; Conservative 61; Mismatches 167; Indels 74; Gaps 17;

QY 117 EFKLEF-EYQTVNVSVKLNSFIPEEG---SDDKPEPFLCSKPKRRORGRNNTRL 172
Db 34 EGIKIFRPYGTSDLDIGKMLPGSGGATADSKFEKFAKNTLADIQYKGDLLYVYV 93
QY 173 KV-----CFLPDSPLTNGTENGITLLDNGNRGLGYPATELAGOFEWTSNI-----PP 222
Db 94 KSKRPTNSLPTLINIAFLDSSERAIKWEYD-----PYTT--TAQWCTATLYKVEPV 144
QY 223 AIAHSLDAGAKVILTSEAVVPATKRLSARSEARSHLLQKRF-----YHSIR 274
Db 145 PPAEGAFKRAVHTLDLSKSGASGRYVSKIKKPTPRSTYEDYKMKMIAKKADKINSFK 204
QY 275 VOPMALEQVMS-----DRDSEDEV--DDVDADFEEDROMLDFVDYVNDKDEK---QFMHL 322
Db 205 -PPKIEFLQSCVLEFVDRSTSDLICGAEPYVEGQYRKYNNSGFVSNDEKRTPOSFSHF 263
QY 323 WNSFYKQKRVAD-----GHISWACEAFSFEKELHRSLSLEFCWRLFLIKLIMHGLVDS 378
Db 264 TYESNHQLLIDIGVG-----DHYTDPQIHYYDGV---GEGIGNIGQKGEFEK 310
QY 379 ATINNCNTILE--NCRNSDPTTTNNNSVDKPS-----DSNTNNNTVDPNDIN 427
Db 311 LDTHKCMACIQLYLMQISNPKSEKSDGTYVPRDLIPPTSEKDNNNNNNN--NNNNNN 368
QY 428 NKNVYDNKDNNSRDKVTK 445
Db 369 NNNNNNNNNNNNNSSISK 386

RESULT 14
US-09-994-485-8
Sequence 8, Application US/09994485
Patent No. US20020142429A1
GENERAL INFORMATION:
APPLICANT: Ryazanov, Alexey G.
Hait, William N.
Pavur, Karen S.
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/994,485
FILING DATE: 27-NO. US20020142429A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostellium discoideum
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-994-485-8

Query Match
Best Local Similarity 4.4%, Score 105; DB 10; Length 732;
Matches 76; Conservative 61; Mismatches 167; Indels 74; Gaps 17;
QY 117 EFKLEF-EYQTVNVSVKLNSFIPEEG---SDDKPEPFLCSKPKRRORGRNNTRL 172
Db 34 EGIKIFRPYGTSDLDIGKMLPGSGGATADSKFEKFAKNTLADIQYKGDLLYVYV 93
QY 173 KV-----CFLPDSPLTNGTENGITLLDNGNRGLGYPATELAGOFEWTSNI-----PP 222
Db 94 KSKRPTNSLPTLINIAFLDSSERAIKWEYD-----PYTT--TAQWCTATLYKVEPV 144
QY 223 AIAHSLDAGAKVILTSEAVVPATKRLSARSEARSHLLQKRF-----YHSIR 274
Db 145 PPAEGAFKRAVHTLDLSKSGASGRYVSKIKKPTPRSTYEDYKMKMIAKKADKINSFK 204
QY 275 VOPMALEQVMS-----DRDSEDEV--DDVDADFEEDROMLDFVDYVNDKDEK---QFMHL 322
Db 205 -PPKIEFLQSCVLEFVDRSTSDLICGAEPYVEGQYRKYNNSGFVSNDEKRTPOSFSHF 263
QY 323 WNSFYKQKRVAD-----GHISWACEAFSFEKELHRSLSLEFCWRLFLIKLIMHGLVDS 378
Db 264 TYESNHQLLIDIGVG-----DHYTDPQIHYYDGV---GEGIGNIGQKGEFEK 310
QY 379 ATINNCNTILE--NCRNSDPTTTNNNSVDKPS-----DSNTNNNTVDPNDIN 427
Db 311 LDTHKCMACIQLYLMQISNPKSEKSDGTYVPRDLIPPTSEKDNNNNNNN--NNNNNN 368
QY 428 NKNVYDNKDNNSRDKVTK 445
Db 369 NNNNNNNNNNNNNSSISK 386

RESULT 15
US-09-858-935B-77
Sequence 77, Application US/09858935B
Publication No. US2003006917A1
GENERAL INFORMATION:
APPLICANT: Dubaquitte, Yves
APPLICANT: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
FILE REFERENCE: P1794R1
CURRENT APPLICATION NUMBER: US/09/858,935B
FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 60/248,985
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 60/204,490
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 77
LENGTH: 97
TYPE: PRT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:35:49 ; Search time 310 Seconds

(without alignments)
925.504 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRKSSPEEVIISDNE.....INNNKNDKNDNRDKYIK 445

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 64473110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: Pending_Patents_AA_Main:*
2: /cgn2_6/ptodata/1/paa/PCITUS.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US088.COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US097.COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US099.COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*
28: /cgn2_6/ptodata/1/paa/US103.COMB.pep:*
29: /cgn2_6/ptodata/1/paa/US104.COMB.pep:*
30: /cgn2_6/ptodata/1/paa/US105.COMB.pep:*
31: /cgn2_6/ptodata/1/paa/US106.COMB.pep:*
32: /cgn2_6/ptodata/1/paa/US107.COMB.pep:*
33: /cgn2_6/ptodata/1/paa/US108.COMB.pep:*
34: /cgn2_6/ptodata/1/paa/US109.COMB.pep:*
35: /cgn2_6/ptodata/1/paa/US110.COMB.pep:*
36: /cgn2_6/ptodata/1/paa/US111.COMB.pep:*
37: /cgn2_6/ptodata/1/paa/US112.COMB.pep:*
38: /cgn2_6/ptodata/1/paa/US113.COMB.pep:*
39: /cgn2_6/ptodata/1/paa/US114.COMB.pep:*
40: /cgn2_6/ptodata/1/paa/US115.COMB.pep:*
41: /cgn2_6/ptodata/1/paa/US116.COMB.pep:*
42: /cgn2_6/ptodata/1/paa/US117.COMB.pep:*
43: /cgn2_6/ptodata/1/paa/US118.COMB.pep:*
44: /cgn2_6/ptodata/1/paa/US119.COMB.pep:*
45: /cgn2_6/ptodata/1/paa/US120.COMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2378	100.0	445	22	US-09-890-220-2
2	2285	96.1	440	22	US-09-890-220-5
3	1923	80.9	498	19	US-09-513-996A-69039
4	1796	75.5	367	19	US-09-513-996A-69040
5	1635	68.8	339	19	US-09-513-996A-69041
6	1153.5	48.5	428	26	US-10-219-999-37133

7	1153.5	48.5	428	27	US-60-324-109-23364	Sequence 23364, A
8	923	38.8	611	17	US-09-339-947A-1	Sequence 1, Appl
9	906.5	38.1	604	17	US-09-339-947A-8	Sequence 8, Appl
10	833	35.0	481	26	US-10-219-999-48099	Sequence 48099, A
11	833	35.0	481	27	US-60-312-544-10052	Sequence 10052, A
12	654	27.5	813	17	US-09-398-237-2	Sequence 2, Appl
13	549	23.1	107	22	US-09-890-220-8	Sequence 8, Appl
14	510.5	21.5	692	25	US-10-177-478-6	Sequence 6, Appl
15	451.5	19.0	186	22	US-09-880-220-14	Sequence 14, Appl
16	423	17.8	743	21	US-09-708-427-25830	Sequence 25830, A
17	423	17.8	765	21	US-09-708-427-25829	Sequence 25829, A
18	423	17.8	851	21	US-09-708-427-25828	Sequence 25828, A
19	368.5	15.5	108	22	US-09-890-220-12	Sequence 12, Appl
20	327	13.8	63	19	US-09-513-996A-69043	Sequence 69043, A
21	324	13.6	145	19	US-09-513-996A-79541	Sequence 79541, A
22	324	13.6	174	19	US-09-513-996A-79540	Sequence 79540, A
23	322.5	13.6	623	21	US-09-708-427-25832	Sequence 25832, A
24	311	13.1	154	22	US-09-890-220-10	Sequence 10, Appl
25	300	12.6	582	21	US-09-708-427-19473	Sequence 19473, A
26	266	11.2	295	21	US-09-708-427-25824	Sequence 25824, A
27	249	10.5	366	26	US-10-219-999-61010	Sequence 61010, A
28	249	10.5	366	27	US-60-324-109-29363	Sequence 29363, A
29	239	10.1	289	21	US-09-708-427-25825	Sequence 25825, A
30	225	9.5	388	1	US-09-764-864-931	Sequence 931, App
31	225	9.5	388	21	PCT-US01-01341-931	Sequence 931, App
32	225	9.5	388	24	US-10-080-129-931	Sequence 931, App
33	224.5	9.4	302	21	US-09-708-427-19475	Sequence 19475, A
34	224.5	9.4	427	21	US-09-708-427-19474	Sequence 19474, A
35	221	9.3	367	22	US-09-890-220-54	Sequence 54, Appl
36	221	9.3	739	1	PCT-US01-17936-5	Sequence 54, Appl
37	221	9.3	739	22	US-09-874-162A-5	Sequence 5, Appl
38	221	9.3	776	1	PCT-US01-17936-8	Sequence 8, Appl
39	221	9.3	776	22	US-09-874-162A-8	Sequence 8, Appl
40	221	9.3	803	22	US-09-890-220-16	Sequence 16, Appl
41	219	9.2	835	16	US-09-270-767-60845	Sequence 60845, A
42	219	9.2	835	16	US-09-270-8498-191180	Sequence 191180, A
43	219	9.2	947	16	US-09-270-767-45347	Sequence 45347, A
44	219	9.2	947	27	US-60-167-324-1120	Sequence 1120, Ap
45	219	9.2	955	20	US-09-614-150-41697	Sequence 41697, A

ALIGNMENTS

```
RESULT 1
US-09-890-220-2
; Sequence 2, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
; TITLE OF INVENTION: vermalisation gene VRN2.
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,220
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-890-220-2
```

Query Match 100.0%; Score 2378; DB 22; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.4e-235;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MCRONCRKSSPEEVIISDNEILLYCKEVLNIFHLRSLGNPSFLPCLNKTGAKRR 60
|||||

Fri Jun 20 09:03:40 2003

us-09-890-220-2.ram

Page 2

```
Db 1 MCRONRAKSSPEEVISTDBENLLIYCKPVRLYNIFHLRSIGNPSFLPRCLNKTGAKRR 60
QY 61 KSRSTGMVYVFNKDCNNLTOKTEVREDCSPSCMLGSGFKLOPHLMSHSDPEPEFKL 120
Db 61 KSRSTGMVYVFNKDCNNLTOKTEVREDCSPSCMLGSGFKLOPHLMSHSDPEPEFKL 120
QY 121 FEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRORGRNNTRELKVCPLD 180
Db 121 FEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRORGRNNTRELKVCPLD 180
QY 121 FEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRORGRNNTRELKVCPLD 180
Db 121 FEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRORGRNNTRELKVCPLD 180
QY 181 SPSTLNGTEGTLTLDNGNRGLGYPEATELAGOPEMTSNIPALHSSLDGAKVILTSE 240
Db 181 SPSTLNGTEGTLTLDNGNRGLGYPEATELAGOPEMTSNIPALHSSLDGAKVILTSE 240
QY 241 AVVPATKTRKLSAERSEANSHLLLOKROFYHSHRVOPMALDOVMSDRDSEDDVDVADF 300
Db 241 AVVPATKTRKLSAERSEANSHLLLOKROFYHSHRVOPMALDOVMSDRDSEDDVDVADF 300
QY 301 EDROMLDDFVYVNDKDEKOPMHLMSFVRKQRYIADGHI SWACEAFSREYEKELHRYSLF 360
Db 301 EDROMLDDFVYVNDKDEKOPMHLMSFVRKQRYIADGHI SWACEAFSREYEKELHRYSLF 360
QY 361 MCWRFLIKLNMHGLVDSATINNCTTLENCNNSDPTTNNNNNSVDRPSDNTNNNTIV 420
Db 361 MCWRFLIKLNMHGLVDSATINNCTTLENCNNSDPTTNNNNNSVDRPSDNTNNNTIV 420
QY 421 DHPNDINNNKNNVNDKNNSDRYK 445
Db 421 DHPNDINNNKNNVNDKNNSDRYK 445
```

RESULT 2

```
US-09-890-220-5
; Sequence 5, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
; FILE REFERENCE: Newbury
; CURRENT APPLICATION NUMBER: US/09/890,220
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/G800/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-890-220-5
```

Query Match 96.1%; Score 2285; DB 22; Length 440;

Best Local Similarity 96.8%; Pred. No. 5,6e-226; Matches 428; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```
QY 1 MCRONRAKSSPEEVISTDBENLLIYCKPVRLYNIFHLRSIGNPSFLPRCLNKTGAKRR 60
Db 1 MCRONRAKSSPEEVISTDBENLLIYCKPVRLYNIFHLRSIGNPSFLPRCLNKTGAKRR 60
QY 61 KSRSTGMVYVFNKDCNNLTOKTEVREDCSPSCMLGSGFKLOPHLMSHSDPEPEFKL 120
Db 61 KSRSTGMVYVFNKDCNNLTOKTEVREDCSPSCMLGSGFKLOPHLMSHSDPEPEFKL 120
QY 121 FEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRORGRNNTRELKVCPLD 180
Db 121 FEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRORGRNNTRELKVCPLD 180
QY 181 SPSTLNGTEGTLTLDNGNRGLGYPEATELAGOPEMTSNIPALHSSLDGAKVILTSE 240
Db 181 SPSTLNGTEGTLTLDNGNRGLGYPEATELAGOPEMTSNIPALHSSLDGAKVILTSE 240
```

```
QY 241 AVVPATKTRKLSAERSEANSHLLLOKROFYHSHRVOPMALDOVMSDRDSEDDVDVADF 300
Db 241 AVVPATKTRKLSAERSEANSHLLLOKROFYHSHRVOPMALDOVMSDRDSEDDVDVADF 300
QY 301 EDROMLDDFVYVNDKDEKOPMHLMSFVRKQRYIADGHI SWACEAFSREYEKELHRYSLF 360
Db 301 EDROMLDDFVYVNDKDEKOPMHLMSFVRKQRYIADGHI SWACEAFSREYEKELHRYSLF 360
QY 361 MCWRFLIKLNMHGLVDSATINNCTTLENCNNSDPTTNNNNNSVDRPSDNTNNNTIV 420
Db 361 MCWRFLIKLNMHGLVDSATINNCTTLENCNNSDPTTNNNNNSVDRPSDNTNNNTIV 420
QY 421 DHPNDINNNKNNVNDKNNSDRYK 442
Db 421 DHPNDINNNKNNVNDKNNSDRYK 440
```

RESULT 3

```
US-09-513-996A-69039
; Sequence 69039, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69039
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..498
; OTHER INFORMATION: any n or xaa = unknown
; OTHER INFORMATION: Location 1..498 / Ceres Seq. ID 2185973
US-09-513-996A-69039
```

Query Match 80.9%; Score 1923; DB 19; Length 498;

Best Local Similarity 89.4%; Pred. No. 1,6e-188; Matches 370; Conservative 7; Mismatches 7; Indels 34; Gaps 3;

```
QY 43 PSFLPRCLNKTGAKRRKRSSTGMVYVFNKDCNNLTOKTEVREDCSPSCMLGSGFK 102
Db 108 PSFLPRCLNKTGAKRRKRSSTGMVYVFNKDCNNLTOKTEVREDCSPSCMLGSGFK 167
QY 103 LOFHNSHDLFEPEFKLFEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRRO 162
Db 168 LOFHNSHDLFEPEFKLFEEOYVNVSVKLNSTIFEEEGSDDDKFEPSLCSKPRKRRO 204
QY 163 RGRNNTRELKVCPLDPSLNGTEGTLTLDNGNRGLGYPEATELAGOPEMTSNIP 222
Db 205 RGRNNTRELKVCPLDPSLNGTEGTLTLDNGNRGLGYPEATELAGOPEMTSNIP 264
QY 223 ALHSSLDGAKVILTSEAVVPATKTRKLSAERSEANSHLLLOKROFYH 271
Db 265 ALHSSLDGAKVILTSEAVVPATKTRKLSAERSEANSHLLLOKROFYH 324
QY 272 SHRVOPALDOVMSDRDSEDDVDVADFEDROMLDDFVYVNDKDEKOPMHLMSFVRKOR 331
Db 325 SHRVOPALDOVMSDRDSEDDVDVADFEDROMLDDFVYVNDKDEKOPMHLMSFVRKOR 384
QY 332 VIADGHISWACEAFSREYEKELHRYSLFMCWRFLIKLNMHGLVDSATINNCTTLENC 391
Db 385 VIADGHISWACEAFSREYEKELHRYSLFMCWRFLIKLNMHGLVDSATINNCTTLENC 444
QY 392 RNSSDPTTNNNNNSVDRPSDNTNNNTIVDHPNDINNNKNNVNDKNNSDRYK 445
Db 445 RNSSDPTTNNNNNSVDRPSDNTNNNTIVDHPNDINNNKNNVNDKNNSDRYK 498
```

RESULT 4

```

US-09-513-996A-69040
; Sequence 69040, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THEREBY
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69040
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..367
; OTHER INFORMATION: any n or xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..367 / Ceres Seq. ID 2185974
US-09-513-996A-69040

```

```

Query Match          75.5%; Score 1796; DB 19; Length 367;
Best Local Similarity 88.7%; Pred. No. 1.3e-175;
Matches 346; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

```

```

QY 67 MVENYKDCNNTLOKTEVREDSCPCSMCGSEKGLQFHLNNSHDLFEFEKLEEEYQT 126
DB 1 MVEFYKDCNNTLOKTEVREDSCPCSMCGSEKGLQFHLNNSHDLFEFEKLEEEYQT 56
QY 127 VNVSKLNSFTFEESGSDDDKFEFPRSLCSKPRKRQRGRNTRRLKVCFLPLDPSLTN 186
DB 57 VNPFTMCS-----SKPRKRQRGRNTRRLKVCFLPLDPSLTN 97
QY 187 GTENGITLLDNGNRGLGYPEATELAGOPEMTSNIPPAIAHSSLDGAKVILTSEAVPAT 246
DB 98 GTENGITLLDNGNRGLGYPEATELAGOPEMTSNIPPAIAHSSLDGAKVILTSEAVPAT 157
QY 247 KTRKLSAERSAR-----SHLLQKQFYHSHRVOPMALEQVMSDRSDSEYVD 295
DB 158 KTRKLSAERSARSTAILTERISSSHLLQKQFYHSHRVOPMALEQVMSDRSDSEYVD 217
QY 296 DVADPEDROMLDDFYVNDKDEKOFHMLNSFVRKORVADGILISACAFSFEYKEELHR 355
DB 218 DVADPEDROMLDDFYVNDKDEKOFHMLNSFVRKORVADGILISACAFSFEYKEELHR 277
QY 356 YSSLEFCWRLFLIKLNMHGLVDSATINNNTILNENCRNSSDTTNNNSVDPSDNTN 415
DB 278 YSSLEFCWRLFLIKLNMHGLVDSATINNNTILNENCRNSSDTTNNNSVDPSDNTN 337
QY 416 NNNIYDHPNDINKNNVNDKNNNSRDYIK 445
DB 338 NNNIYSHPDINKNNVNDKNNNSRDYIK 367

```

RESULT 5

```

US-09-513-996A-69041
; Sequence 69041, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THEREBY
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69041
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:

```

```

; NAME/KEY: UNSURE
; LOCATION: 1..339
; OTHER INFORMATION: any n or xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..339 / Ceres Seq. ID 2185975
US-09-513-996A-69041

```

```

Query Match          68.8%; Score 1635; DB 19; Length 339;
Best Local Similarity 87.8%; Pred. No. 4.6e-159;
Matches 318; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

```

```

QY 95 MCGSEKGLQFHLNNSHDLFEFEKLEEEYQTVNVSKLNSFTFEESGSDDDKFEFSLC 154
DB 1 MCGSEKGLQFHLNNSHDLFEFEKLEEEYQTVNVSKLNSFTFEESGSDDDKFEFSLC 37
QY 155 SKPRKRQRGRNTRRLKVCFLPLDPSLTNLTENGITLLDNGNRGLGYPEATELAGO 214
DB 38 SKPRKRQRGRNTRRLKVCFLPLDPSLTNLTENGITLLDNGNRGLGYPEATELAGO 97
QY 215 EMTSNIPPAIAHSSLDGAKVILTSEAVPATKTRKLSAERSAR-----SHLL 263
DB 98 EMTSNIPPAIAHSSLDGAKVILTSEAVPATKTRKLSAERSARSTAILTERISSSHLL 157
QY 264 LQKQFYHSHRVOPMALEQVMSDRSDSEYVDVADFDERQMLDDFYVNDKDEKOFHML 323
DB 158 LQKQFYHSHRVOPMALEQVMSDRSDSEYVDVADFDERQMLDDFYVNDKDEKOFHML 217
QY 334 NSFVRKQVADGILISACAFSFEYKEELHRYSSLEFCWRLFLIKLNMHGLVDSATINN 383
DB 218 NSFVRKQVADGILISACAFSFEYKEELHRYSSLEFCWRLFLIKLNMHGLVDSATINN 277
QY 384 CNTILENCRNSSDTTNNNSVDPSDNTNNNIYHPNDINKNNVNDKNNSRKV 443
DB 278 CNTILENCRNSSDTTNNNSVDPSDNTNNNIYHPNDINKNNVNDKNNSRKV 337
QY 444 IK 445
DB 338 IK 339

```

RESULT 6

```

US-10-219-999-37133
; Sequence 37133, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219, 999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 37133
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Glycine max
US-10-219-999-37133

```

```

Query Match          48.5%; Score 1153.5; DB 26; Length 428;
Best Local Similarity 54.0%; Pred. No. 3.3e-109;
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;

```

```

QY 1 MCRONCRACKSPPEEYISDENILLYCKPVRLYNIFHLRSLGNSFLPCLNKKYKAKRR 60
DB 1 MCRONSPVHHAGEEIEADESLIYCKVELEYLNILYRALQNPFLRCLRKIKIASKR 60

```

[illegible]

```

RESULT 7
US-60-324-109-23364
Sequence 23364, Application US/60324109
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkle, Gregory J.
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua
TITLE OF INVENTION: DNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-101521261B
CURRENT APPLICATION NUMBER: US/60/324, 109
CURRENT FILING DATE: 2001-09-21
NUMBER OF SEQ ID NOS: 33196
SEQ ID NO 23364
LENGTH: 428
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
US-60-324-109-23364

```

Query Match	48.5%	Score 1153.5	Ds 27	Length 428
Best Local Similarity	54.0%	Pred. No. 3.3e-109		
Matches	235	Conservative	57	Mismatches 102; Indels 41; Gaps 8
QY	1	MCRONRAKSESEEEYSTDENLILLYCPVRRLYNLFHLRSNGNSFPLPCLANKIGAKRRR	60	
Ds	1	MORONSPVHMGEEIEADESLLILLYCPVELYNLYLRALQNSFPLRCLRYKIRRSRR	60	
QY	61	KSRNGGVVFNKDCNNLTQKTEVEKDCSPFCSMGLGSPKGFPLNLSHLLFEFEFL	120	
Ds	61	RLRA-GIVFNFNRDYNILRLKTEVETEDFSCFCLMCGNKGKRFHLCSHLLFNFEVW	119	
QY	121	FEFYQVNVSVLNSFFIIFEEGSDSDD--DKFEPSLSCSKRRKROGSGNNTRRLKVCFL	177	
Ds	120	TEDQYAVNVSVKID--ILRSENVADGVIPQSQFFFCRSRRKRRRDSVQIKRINVFEL	177	
QY	178	PLDPSLANGTENGITLLND-----GNRGGEYEAREL	210	
Ds	178	ELDSP--EGHNGFLQKDDILSCKEENVSRTRSSEKIPPSGRNDGGFGFGDHGTMDN	234	
QY	211	AGQFEKTSNPP--ALAHSSLDAG-AKVILTSEAVVPATKRLLSAERSSEARSHLLQK	266	

D3	233	LEHVESFNTPGVSTAMPSSVDPBCKNSICKSDPALPA - KTKKLSMDRSDSRMRMLDQ	2393
QY	267	KQYVHSHRQYPMALQDYMSDRDSEFQYRDDVADPFRRDMLADPDYVUKKDEQFQYHNLNSP	3265
D3	294	RLFHSHRQYPMALQDYVLSDRDSEBVDVDDIADLEDRMLDDEFDYVSKDEQKQYHNLNSP	3533
QY	327	VRRQRYTADGHSMAQCEAFSEYKEKYLHRSLSLFWQWRLFTKTLNNGLYVDSATLNNCPT	3866
D3	354	MRRQRYTADGHSVMAQCEAFSKLQKELLSPALFWQWRLFWIKLNNGLDQACTMNNCST	4133
QY	387	ILENCNNSDPTTNTN	401
D3	414	VLDSTYNNEGSGTKRN	428

```

RESULT 8
US-09-339-947A-1
Sequence 1, Application US/09393947A
GENERAL INFORMATION:
APPLICANT: YOSHIDA, Nobumasa
APPLICANT: YAMAI, Yukihito
APPLICANT: KATO, Yoshihiro
APPLICANT: HIRATSUKA, Junzo
APPLICANT: TAKAHASHI, Shigeru
APPLICANT: MIWA, Tatsushi
TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
FILE REFERENCE: 032735-001
CURRENT APPLICATION NUMBER: US/09/339,947A
CURRENT FILING DATE: 1999-05-25
PRIORITY APPLICATION NUMBER: JP 1998-180065
PRIORITY FILING DATE: 1998-06-26
PRIORITY APPLICATION NUMBER: JP 1999-179043
PRIORITY FILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 611
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: ZN_FING
LOCATION: (306)..(327)
US-09-339-947A-1

```

[illegible]

QY 145 -DDKEPESLCKP-RKRRGRGNTRRLKVCFLPLDPSLTNG-----TENGITLLN 196
 DB 361 VPKQQTFFESSKRRRRRQKQVSSRQ-----CPHGLGCEVLDKTDHDAHSVS 411
 QY 197 DGNR---GIGYPEATELAGOFEPMTSNIP---PALAHSSLDAGAVIITSEAVPATKTR 249
 DB 412 EKSRIIPGKH---ERIGGAESGGRVPGTSPALVOQSGDDPYQVSIAGSTMLOFAKTR 467
 QY 250 KLSARSEARSHLLLOKROFYHSHRVOPALRQVMSDRDSEDEVDVADFEEDROMLDF 309
 DB 468 KSTIRSDRNSLLOKROFHSRPAQPALRQVMSDRDSEDEVDVADFEEDROMLDF 527
 QY 310 VDYNDEKQFMHLMNSFVRKQVVIADGHISMACEAFSREYKELHRYSLFMCWRLFLIK 369
 DB 528 VDVTEDEKQMMHNSFVRKQVVIADGHISMACEAFSRLHGPIMVTRPHLMCMRYEMWK 587
 QY 370 LAMHGLVDSATINNCNTILE 389
 DB 588 LAMHGLVDSATINNCNTILE 607

RESULT 9
 US-09-339-947A-8
 ; Sequence 8, Application US/09339947A
 ; GENERAL INFORMATION:
 ; APPLICANT: YOSHIDA, Nobumasa
 ; APPLICANT: YANAI, Yukihito
 ; APPLICANT: KATO, Yoshihiro
 ; APPLICANT: HIRAKISUKA, Junzo
 ; APPLICANT: TAKAHASHI, Shigeru
 ; APPLICANT: MIWA, Tatsushi
 ; TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
 ; FILE REFERENCE: 032735-001
 ; CURRENT APPLICATION NUMBER: US/09/339, 947A
 ; CURRENT FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: JP 1998-180065
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: JP 1999-179043
 ; PRIOR FILING DATE: 1999-06-24
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 604
 ; TYPE: PRF
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: 2N_FING
 ; LOCATION: (310)..(335)
 US-09-339-947A-8

Query Match 38.1%; Score 906.5; DB 17; Length 604;
 Best Local Similarity 33.3%; Pred. No. 1.8e-83;
 Matches 209; Conservative 59; Mismatches 109; Indels 251; Gaps 7;
 QY 1 MCRORCAKSPDEEYISTDENLTIYCKPRLYNIFRLSGNPSFLPCLNKGKRRR 60
 DB 1 MCRORPARLSPDEOLAEESPALCYKPELVETLNIORRSIKNPAFLORCLYKHARRKK 60
 QY 61 KSRST----- 65
 DB 61 RSLITISLGGTNKELRAQNIPIPLVYLARPTNNVSLSEHSPIYRSRACILTSFHEFGN 120
 QY 66 ----- 65
 DB 121 KDYTEATFVIPDKMLATSRACSLNIIISGRABQTFDDNNGSGNHVSGSTLOKLEGKC 180
 QY 66 ----- 65
 DB 181 FNGKIPIDLLASLGNVSLIGHVEMSVTEAMPSFLPEKFLDSDCLTFCSSQKVDAT 240
 QY 66 -----GMVFNKDCNNTLQ 80

DB 241 GSFOLVISAQEAQAKDMSSEPSYVSYNDVPSSLTHIIRLGNVLFNKYNNMTQ 300
 QY 81 KTEVEDDCSCPCSMCGSFKLOHLSHSHLFEEFEFLFREYQVNVSVKNSFIPE- 139
 DB 301 KTEVEDDCSCPCPLVPCGSEFKLGCHLNSHLEFEFEIISECAVNVSLTDSRTKL 360
 QY 140 -EEGSDDDKFEFEPSLCKSPRRK-RORGNNTRRLKVCFLPLDSP-SLTNGEITLLN 196
 DB 361 LAGEV-DPRHQFFSRSRKRRKKRYEISDKLRHVPHLYVDSGSPDQAQSGEDDVORE 419
 QY 197 DGNRGLGYPEATELAGOFEPMTSNIPALAHSSLDAGAVY---ITSEAVPATKTRKLSA 253
 DB 420 NGS-----SVAHASVDPAANSIHGSLNLSAPTVLQFGTKRLKSV 456
 QY 254 ERSSEARSHLLLOKROFYHSHRVOPALRQVMSDRDSEDEVDVADFEEDROMLDFPVYN 313
 DB 457 ERADPRNROLLOKROFHSRPAQPALRQVMSDRDSEDEVDVADFEEDROMLDFVDT 516
 QY 314 KDEKQFMHLMNSFVRKQVVIADGHISMACEAFSREYKELHRYSLFMCWRLFLIKLNMH 373
 DB 517 KDEKLMHMMNSFVRKQVVIADGHISMACEAFSFGQELVONPALMCMRYEMWKLMH 576
 QY 374 GLVDSATINNCNTILENCSSDPTTTN 401
 DB 577 SLIDARAMNACNTILEGYNSSDPKK 604

RESULT 10
 US-10-219-999-48099
 ; Sequence 48099, Application US/102199999
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Stein, Joshua
 ; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-10(52726)C
 ; CURRENT APPLICATION NUMBER: US/10/219, 999
 ; CURRENT FILING DATE: 2002-08-15
 ; PRIOR APPLICATION NUMBER: US 60/324,109
 ; PRIOR FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/312,544
 ; PRIOR FILING DATE: 2001-08-15
 ; NUMBER OF SEQ ID NOS: 63520
 ; SEQ ID NO 48099
 ; LENGTH: 481
 ; TYPE: PRF
 ; ORGANISM: Zea mays
 US-10-219-999-48099

Query Match 35.0%; Score 833; DB 26; Length 481;
 Best Local Similarity 48.6%; Pred. No. 4.9e-76;
 Matches 180; Conservative 45; Mismatches 83; Indels 62; Gaps 8;
 QY 43 PSFLPCLNKGKRRKRSRSTGMVYVYKCCNNTLQTEVREDCSCPCSMCGSFNG 102
 DB 153 PSFLPEITRLR-----AGNVLFYKYYNNMTMORTETEDPSCPCLVPCGCSFNG 201
 QY 103 LOFHLSHSHDLFEFEFKLFEFYQTVNVSVKLSF--IPEEGSDDDKFPFSLCKSPRRK 160
 DB 202 LGCHLNSHDLHYELMISEEQAVNVSLKADAMKTEFAVEGY-DPRHQTFSCSFKRR 260
 QY 161 RORGNNTRRLKVCFLPLDPSLTNGT-----ENGITLLDGNRGLGYP 205
 DB 261 R-----RLETTAEK--FRHVHPHIMESGSHEDAQAGEDDYVORENGLSVAN----- 305
 QY 206 EATELAGOFEPMTSNIPALAHSSLDAGAVIITSEAVPATKTRKLSARSRARSHLLQ 265
 DB 306 -----TSVDSOPVHG-----NLSPTVLOFGKSRLKLSARSDPRNROLLO 347
 QY 266 KROFYHSHRVOPALRQVMSDRDSEDEVDVADFEEDROMLDFVDYNDEKQFMHLMNS 325

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapm

Page 6

Db	348	KRF ⁺ FFSHRAQPP ⁺ LEQVY ⁺ SDSDSEDEDD ⁺ DADE ⁺ RRMLD ⁺ DE ⁺ VDVY ⁺ KDKEL ⁺ IMHWN ⁺ S	4078
QY	326	FVRKRYADGHTSMACEAFSR ⁺ FYE ⁺ KE ⁺ LHRS ⁺ SL ⁺ FWCRL ⁺ L ⁺ IKLMNG ⁺ GLVDSAT ⁺ INN ⁺ CN	3859
Db	408	FVRKRYADGHT ⁺ PMACEAFSR ⁺ CHQOQLV ⁺ QNPALL ⁺ MG ⁺ RFF ⁺ MIKIMHNS ⁺ L ⁺ DARANAC ⁺ N	4678
QY	386	TILENCRNSS	395
Db	468	TILQGYDQGS	477

```

RESULT 11
US-60-312-544-10052
Sequence 10052, Application US/60312544
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkley, Gregory J.
APPLICANT: Kovallic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua
TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-10(52726)A
CURRENT APPLICATION NUMBER: US/60/312,544
CURRENT FILING DATE: 2001-08-15
NUMBER OF SEQ ID NOS: 10730
SEQ ID NO 10052
LENGTH: 481
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3206-226-G12_F11
US-60-312-544-10052

```

Query Match	35.0%;	Score 833;	DB 27;	Length 481;
Best Local Similarity	48.6%;	Pred. No. 4.9e-76;		
Matches 180;	Conservative 45;	Mismatches 83;	Indels 62;	Gaps 8

```

QY 4 PSTPLPCLYKTKGAKRRKRKRSTGVVYVNYDCNNLTÖKTVEBDCSPCSCMYCSFK 102
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 153 PSLPLPEIIRLR-----AGVULENTKYNYNNMQKTVEYBDESPCPLVPCSGFK 201
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 103 LOEHLNNSHDLFEFEFKTLEFEYQVYVNVTKLNSP--TFEEGSDDDKFEPEPSCSKRKR 160
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 202 LCGHLNNSHDLTHYELMTISEEQAVNVSLKDAKTEFVAGV-DPRHÖTFYCSRKRR 260
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 161 RQNGGNNNRRLKAYGELPLDPSYLNGF-----ENGTLLDNCRGLGYP 205
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 261 R-----RLLETAKR--FRHYVHYIMSSGHEDAQOAGSEDDYQÖRENGSLVAV----- 305
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 206 EAEELLAGQFEMTSNIPRALHNSLDGAKVILTSEAVYPAIKTKRLSAEFSSEARSHILLQ 265
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 306 -----TSVDPDQVHGS-----NLSPEYVQÖFGSKSKLSAENSDDPRNROLL 347
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 266 KRQFYSHHAYQMALEQVMSDRDSEDDVDVADPEDBRÖMIDFVDVYNNKDEKOPHILWNS 325
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 348 KRQFESHSHAQMPLEQVLTSDHSDSEDDVDVADPEDBRMDLDVDTYKÖEKILMHWNNS 407
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 326 FYRKQVYLDGHTISWACAFSFRFEYKEKILHRYSSLFEWCVRLLRLKLNHGLVDSATTINCN 385
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 408 FYRKQVYLDGHTILPACAFSFRKCHQÖQLVONPALLMGMRFFMIKLNHNSLLDARAVNACN 467
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 386 TILLENCRNS 395
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 468 TILQGYODDS 477
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

RESULT 12
US-09-398-237-2
; Sequence 2, Application US/09398237
; GENERAL INFORMATION:
APPLICANT: Commonwealth Scientific and Industrial Research Organisation

```

1  TITLE OF INVENTION: Novel method of regulating seed development in plants and
2  FILE REFERENCE: P:\opei\mro\autseed.pct
3  CURRENT APPLICATION NUMBER: US/09/398,237
4  CURRENT FILING DATE: 1999-09-20
5  EARLIER APPLICATION NUMBER: US 60/101184
6  EARLIER FILING DATE: 1998-09-21
7  EARLIER APPLICATION NUMBER: AU PP6063
8  EARLIER FILING DATE: 1998-09-22
9  EARLIER APPLICATION NUMBER: AU PP6062
10 EARLIER FILING DATE: 1998-09-22
11 EARLIER APPLICATION NUMBER: AU PP6061
12 EARLIER FILING DATE: 1998-09-22
13 EARLIER APPLICATION NUMBER: AU PQ1345
14 EARLIER FILING DATE: 1999-07-01
15 EARLIER APPLICATION NUMBER: AU PQ1346
16 EARLIER FILING DATE: 1999-07-01
17 NUMBER OF SEQ ID NOS: 213
18 SOFTWARE: PatentIn Ver. 2.0
19 SEQ ID NO: 2
20 LENGTH: 813
21 TYPE: PRT
22 ORGANISM: Arabidopsis thaliana
23 US-09-398-237-2

```

Query Match	27.58;	Score 654;	DB 17;	Length 813;
Best Local Similarity	24.28;	Pred. No. 3.3e-57;		
Matches 184;	Conservative 65;	Mismatches 147;	Indels 364;	Gaps 9;

[illegible]

Qy 319 FMHLMNSFVRKORVIADGHISMACEAFSEFEKELEHRYSSLFMCWRLFLIKLNMGLVDS 378
 Db 716 YMLNINIFVRKORVIADGHIVPACEEFAKLHEEKKNSSPFDMWRMRRIKIMNGLICA 775
 Qy 379 AFINNCNTILENCRRNSDPTTTNNNSVDRPSDSNTNNN 418
 Db 776 KTFHCKTITL-----LSNDEAGQFTSGSAANANN 805

RESULT 13

US-09-890-220-8
 ; Sequence 8, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 107
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-8

Query Match 23.1%; Score 549; DB 22; Length 107;
 Best Local Similarity 99.0%; Pred. No. 9.5e-48;
 Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCRONCRASSPEEYISTDENLILYCKPVRLYNIFHLISLGNPSFLPCLNKYKIGAKRRR 60
 Db 1 MCRONCRASSPEEYISTDENLILYCKPVRLYNIFHLISLGNPSFLPCLNKYKIGAKRRR 60

Qy 61 KSRSTGAVYFNKDCNNLTQKTEVEDSCSPFCSMGCSFK 101
 Db 61 KSRSTGAVYFNKDCNNLTQKTEVEDSCSPFCSMGCSFK 101

RESULT 14

US-10-177-478-6
 ; Sequence 6, Application US/10177478
 ; GENERAL INFORMATION:
 ; APPLICANT: Dang, Van-Dinh
 ; APPLICANT: Okamoto, Jack
 ; TITLE OF INVENTION: Chimeric Histone Acetyltransferase
 ; FILE REFERENCE: Polypeptides
 ; CURRENT APPLICATION NUMBER: US/10/177,478
 ; PRIOR FILING DATE: 2002-06-21
 ; PRIOR APPLICATION NUMBER: U.S. 60/300,135
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 692
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-177-478-6

Query Match 21.5%; Score 510.5; DB 25; Length 692;
 Best Local Similarity 22.3%; Pred. No. 1.7e-42;
 Matches 154; Conservative 53; Mismatches 121; Indels 363; Gaps 8;

Qy 81 KTEVEDSCSPFCSMGCSFKLOFHLNNSHDLFEFEKLEEXYO-TVNVSVKLNSEIF 138
 Db 4 KAEVVENSCPFCLIPCGHGGLQLHLKSSHDAPKFEFEYRAEKDHGPEVDVSKSDTLKF 63

Qy 139 E--EEGDDDKFEPELSCKRP-KRROGRNTRRLKVCFLPLDSPSLTNGTENGITLL 195
 Db 64 GYLKDDVGNPOLSPILTFCSKRNQRÖDDSNNAKLVLMELDLDPSTENDSTHV 123
 Qy 196 NDGN----- 199
 Db 124 NDDNVSPRAHSSKISDILTTQOLAIESSEPKVPHVNDGNVSPRAHSSAKNEST 183
 Qy 200 ----- 199

Qy 184 HVNDDDDVSPRAHSSKNESTHYNEDNISPPRAHSSKNESTHNDDEDVSPPTRS 243
 Qy 200 -----RG----- 201

Qy 244 SKETSIDILTTQPAIVESEPKVRKGRKÖLYAKRYARETÖPAIAESSEPKYLHVND 303
 Qy 202 --LGYPEATELAGOEMTSNIPPAIAHSS----- 228

Qy 304 NVSSPPEAHSLKASDILTTQPAIAESSEPKVPHVNDENVSTPRAHSSKKNSTKRV 363
 Qy 229 ----- 228

Qy 364 DNVSPPKTRSSKTSIDILTTQPTIAESSEPKVPHVNDNVSTPRAHSSKKNSTKRN 423
 Qy 229 -----LDAGAVIILSEAVV----- 244

Qy 424 DNDIPSPKTRSSKTSNLTFTQPAIAESSEPKVPHVNDKYSTPRAHSSKKNSTHKK 483
 Qy 245 -----ATTKRLSAERSEARSHLLÖR 267

Qy 484 DDMASLPKTRSSKTSIDILTTQPAIAEPSEPKTVRSRKELHAERCEAKRLRLGR 543
 Qy 268 QFYSHRVQPALEOVMSDRSEDEVDVADFEEDÖMLDEVDVNDKDEKÖFHLNNSFV 327

Qy 544 QFYHSÖTQÖPMTFEOYMSNEDSENETDYALDISRLRLERYGVSKERKRYMYLNIIV 603
 Qy 328 KRÖRYIADGHISMACEAFSEFEKELEHRYSSLFMCWRLFLIKLNMGLVDSATINNCNTI 387

Qy 604 KRÖRYIADGHIVPACEEFAKLHEEKKNSSPFDMWRMRRIKIMNGLICATFHKCTTI 663
 Qy 388 LENCRRNSDPTTTNNNSVDRPSDSNTNNN 418

Qy 664 L-----LSNDEAGQFTSGSAANANN 684
 Db

RESULT 15

US-09-890-220-14
 ; Sequence 14, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendall, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 186
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-14

Query Match 19.0%; Score 451.5; DB 22; Length 186;
 Best Local Similarity 50.8%; Pred. No. 2.6e-37;
 Matches 99; Conservative 32; Mismatches 43; Indels 21; Gaps 7;

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapm

Page 8

```
QY 237 LTRAVVPATKTKLSEARSHLL-LQKQFYHSHRVPMALEQVMSDRDSEDEVD- 294
DB 6 LITRAKVP-----AKRSKATSHYLPJHKRQFYHSRTGQPLSLQVMSDRDSEDEVDK 57
QY 295 -DVADEPDRQMLDDFVDVVK-DEKQFMHLMNSFVKQRYADGHIWACEAFSRYEKE 352
DB 58 NDDAAHLESQMLNGSDEDEVEIYAEKFIKLMSFVKQRYADGHIWACEAFSRLHJQE 117
QY 353 LHRYSLEWCMWRLFLIKLNNHGLVDSATJNNCNTILENCRNSSDTTNNNNSYDRPSDS 412
DB 118 LRSNLSLJLCWQFMKQMDYGLDRVTMKNKCNTIYH--NISTNDINNNT-----R 170
QY 413 NTNNNNIVDHPNDIN 427
DB 171 TTDMMDVVD--DDIN 183
```

Search completed: June 13, 2003, 15:43:51
Job time : 313 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein -- protein search, using sw model

Run on: June 13, 2003, 15:30:54 ; Search time 54 Seconds
(without alignments) 2102.183 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378
Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INKNKNVNDKNSRDYIK 445

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1155225 seqs, 255096386 residues

Total number of hits satisfying chosen parameters: 1155225

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1153.5	48.5	428	6	US-10-424-599-254264 Sequence 254264, A
2	1153.5	48.5	435	6	US-10-425-114-46823 Sequence 46823, A
3	833	35.0	481	6	US-10-425-114-54921 Sequence 54921, A
4	654	27.5	813	6	US-10-231-778-2 Sequence 2, Appl1
5	648.5	27.3	640	6	US-10-437-963-193525 Sequence 193525, A
6	617.5	26.0	201	6	US-10-437-963-114092 Sequence 114092, A
7	575.5	24.2	242	6	US-10-424-599-283985 Sequence 283985, A
8	510.5	21.5	692	6	US-10-198-723-6 Sequence 6, Appl1
9	491	20.6	166	6	US-10-424-599-254263 Sequence 254263, A
10	269	11.3	102	6	US-10-424-599-183487 Sequence 183487, A
11	256.5	10.8	116	6	US-10-424-599-216298 Sequence 216298, A
12	255.5	10.7	151	6	US-10-437-963-199908 Sequence 199908, A
13	249	10.5	394	6	US-10-425-114-70148 Sequence 70148, A
14	190	8.0	263	6	US-10-437-963-184120 Sequence 184120, A
15	159	6.7	195	6	US-10-424-599-250979 Sequence 250979, A
16	140.5	5.9	116	6	US-10-424-599-152510 Sequence 152510, A
17	132	5.6	749	6	US-10-369-493-21933 Sequence 21933, A
18	128	5.4	106	6	US-10-437-963-116154 Sequence 116154, A
19	124	5.2	1288	1	PCF-US02-38360-2 Sequence 2, Appl1
20	122	5.1	86	6	US-10-437-963-201278 Sequence 201278, A
21	122	5.1	123	6	US-10-437-963-109658 Sequence 109658, A
22	119	5.0	1094	1	PCF-US01-29288A-216 Sequence 216, App
23	119	5.0	1094	6	US-10-149-310-216 Sequence 216, App
24	118	5.0	249	6	US-10-437-963-135834 Sequence 135834, A
25	115.5	4.9	2184	6	US-10-304-095-6 Sequence 6, Appl1
26	114.5	4.8	879	6	US-10-282-122A-52571 Sequence 52571, A

27	114	4.8	964	1	PCF-US01-29288A-256 Sequence 256, App
28	114	4.8	964	6	US-10-149-310-256 Sequence 256, App
29	113.5	4.8	105	6	US-10-437-963-112367 Sequence 112367, A
30	113.5	4.8	654	5	US-09-134-000C-6645 Sequence 6645, App
31	113.5	4.8	654	5	US-09-134-000C-6645 Sequence 6645, App
32	113.5	4.8	654	5	US-10-434-665-6645 Sequence 6645, App
33	112	4.7	104	6	US-10-437-963-198200 Sequence 198200, A
34	108.5	4.6	905	6	US-10-369-493-21936 Sequence 21936, A
35	108	4.5	336	6	US-10-424-599-223789 Sequence 223789, A
36	108	4.5	336	6	US-10-438-246-32035 Sequence 32035, A
37	108	4.5	337	6	US-10-425-114-39945 Sequence 39945, A
38	107.5	4.5	192	6	US-10-437-963-152724 Sequence 152724, A
39	107.5	4.5	833	6	US-10-369-493-21936 Sequence 21936, A
40	107	4.5	138	6	US-10-437-963-176416 Sequence 176416, A
41	107	4.5	313	6	US-10-425-114-52231 Sequence 52231, A
42	106.5	4.5	833	6	US-10-310-154-480 Sequence 480, App
43	106	4.5	328	6	US-10-437-963-107965 Sequence 107965, A
44	105.5	4.4	787	6	US-10-282-122A-72461 Sequence 72461, A
45	104.5	4.4	97	6	US-10-271-869-77 Sequence 77, Appl

ALIGNMENTS

RESULT 1									
US-10-424-599-254264									
Sequence 254264, Application US/10424599									
GENERAL INFORMATION:									
APPLICANT: La Rosa Thomas J									
APPLICANT: Kovalic David K									
APPLICANT: Zhou Yihua									
APPLICANT: Cao Yongqiang									
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with									
FILE REFERENCE: 38-21(53223)B									
CURRENT APPLICATION NUMBER: US/10/424,599									
CURRENT FILING DATE: 2003-04-28									
NUMBER OF SEQ ID NOS: 285684									
SEQ ID NO 254264									
LENGTH: 428									
TYPE: PRT									
ORGANISM: Glycine max									
FEATURE:									
OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1.pep									
US-10-424-599-254264									
Query Match									
Best Local Similarity 54.0%; Pred. No. 3.9e-100;									
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;									
QY	1	MCRQNCRAKSSPEEVISTDENLLIYCKRYRLNIFHLASLGNPSFLPCINTKIGARRR	60						
DB	1	MCRQNSPVHAAEEIEADESLIYCKPEYELNIIYRLALQNPSTLRCLRKIKIRASRR	60						
QY	61	KSRSGMVFNFYKDKDNTLQKTEVEDSCPCSMICGSKFGLQFHLNSSHDLFFEEKL	120						
DB	61	RRAA-GIYIENRDRYINILRKTEVEDSCPCMLQCGFKRLRHILCSHDLFFEEFY	119						
QY	121	FEETGVNVSVKLSFTFEESGSD--DKPEPFLCSKPRRRRORGRNNTRLKVCFL	177						
DB	120	TEDYQAVNVSVKID--ILRSENVADGVIPQSTFFCSPRRKRRKDSVQIEKRNIVFL	177						
QY	178	PIDSLSLNGTENGTLTND-----GNRGIGYPEATEL	210						
DB	178	EIDSP--EGIHNGFLQKDDILSKGEVNSTSEKIFPSGRNDGKFGFDHGTMDN	234						
QY	211	AGQFEMTSNIP--AIAHSSLDAG-AVYITLSEAVVPATKTRKLSAESEARSHLLQK	266						
DB	235	LEHVSSFNIPQVSIAMPOSSVDPSCSKSICKSDPALPA-KTKKLSMDRSDSRNRLQK	293						
QY	267	ROFYSHRVQPALQVMSDRSEDEVDVADFEEDRQMDLVFVYKNDKQFHLNNSF	326						
DB	294	RLEFHSRVPALBOVLSDRSEDEVDVADFEEDRRLDLDFVYSKDEQOLMLNNSF	353						

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapn

Page 2

OY 327 VRKORVADGHISMACEAFSEYKELHRYSSLFQWCMRLFKIKMNGLDNSATTNNCT 386
DB 354 MKRORVADGHVPMACEAFSEYKELHRYSSLFQWCMRLFKIKMNGLDNSATTNNCT 413
OY 387 ILENCRSSDPTTNN 401
DB 414 VLDSEYRNESGSTRKN 428

RESULT 2
US-10-425-114-46823
Sequence 46823, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 46823
LENGTH: 435
TYPE: PRN
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 701055296_FLI.pep
US-10-425-114-46823

Query Match 48.5%; Score 1153.5; DB 6; Length 435;
Best Local Similarity 54.0%; Pred. No. 46-100;
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;
OY 1 MCRONCRANSSPEVISTDNLITCKPVRLNIFHLISLGNPSFLPCLNYKIGAKRR 60
DB 8 MCRONSPVHAGEEELADESLITCKPVRLNIFHLISLGNPSFLPCLNYKIGAKRR 67
OY 61 KSRSTGVVFNKYKDCNNLTQKTEVEDSCSPCSMLCGSFGLQFHLNSHDLFEFEKL 120
DB 68 RLRA-GIVIFNFRDRIKRLKTEVEDSCSPCSMLCGSFGLQFHLNSHDLFEFEKL 126
OY 121 FEYQTVNVSVKLNSPFEEDSGD--DKPEFSLCKSPKRRQGGRRNTRRLKVCFL 177
DB 127 TEDIQAVNVSVKID--ILRSENVADGVIPDSQTFPFCSPRRRRKDSVOJERTKYNKFL 184
OY 178 PLDSPSLTNGTENGITLND-----GNRGLGYPEATEL 210
DB 185 ELDSP---EGIHNGFLQKDDDLISCKGENVSRKSEKIFPSGRNDGKFGDPHPTNDN 241
OY 211 AGOFEMTSNIP---AIAHSSLDAG-AKVIITSEAVVPATYTRKLSAERSSEARSHLLQ 266
DB 242 LEHVESFPNIPGSIAMPOSSVDPCSKSICKSDPALPA-KTKILMSDRSDSRRLMLQ 300
OY 267 ROYHSHRVPOMALEQVMSDRSEDEVDVADPEDROMLDFDVNKNDEKQFHLNNSF 326
DB 301 RLFEHSHRVPOMALEQVMSDRSEDEVDVADPEDROMLDFDVNKNDEKQFHLNNSF 360
OY 327 VRKORVADGHISMACEAFSEYKELHRYSSLFQWCMRLFKIKMNGLDNSATTNNCT 386
DB 361 MKRORVADGHVPMACEAFSEYKELHRYSSLFQWCMRLFKIKMNGLDNSATTNNCT 420
OY 387 ILENCRSSDPTTNN 401
DB 421 VLDSEYRNESGSTRKN 435

RESULT 3
US-10-425-114-54921

Sequence 54921, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54921
LENGTH: 481
TYPE: PRN
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73226G12_FLI.pep
US-10-425-114-54921

Query Match 35.0%; Score 833; DB 6; Length 481;
Best Local Similarity 48.6%; Pred. No. 86-70;
Matches 180; Conservative 45; Mismatches 83; Indels 62; Gaps 8;
OY 43 PSFLPCLNYKIGAKRRKRSSTGMVFNKYKDCNNLTQKTEVEDSCSPCSMLCGSFGLQ 102
DB 153 PSFLPCLNYKIGAKRRKRSSTGMVFNKYKDCNNLTQKTEVEDSCSPCSMLCGSFGLQ 201
OY 103 LQFHLNSHDLFEFEFEKIFEEYQTVNVSVKLNSF--IFEEDSGDDDKPEFSLCKSPKRR 160
DB 202 LQCHLNSSHDLFHELMVSECCQAVNVSLLADAMKTEFAEGV-DRHQTFSYCSRFKRR 260
OY 161 RQGGRRNTRRLKVCFLPLDSPSLTNGT-----ENGITLNDGNGLGYP 205
DB 261 R---RLTTEAEK--FRVHPHIMESGSHDAQSGEDDYVORENGLSVAN----- 305
OY 206 EATELAGOFEMTSNIPALAHSSLDAGAKVITSEAVVPATYTRKLSAERSSEARSHLLQ 265
DB 306 -----TSVDPSPQPVHGS-----NLSPPTVLOFGKSRKLSAERSSDPRRROLQ 347
OY 266 KROFYHSHRVPOMALEQVMSDRSEDEVDVADPEDROMLDFDVNKNDEKQFHLNNSF 325
DB 348 KROFYHSHRVPOMALEQVMSDRSEDEVDVADPEDROMLDFDVNKNDEKQFHLNNSF 407
OY 326 FVRKORVADGHISMACEAFSEYKELHRYSSLFQWCMRLFKIKMNGLDNSATTNNCT 385
DB 408 FVRKORVADGHVPMACEAFSEYKELHRYSSLFQWCMRLFKIKMNGLDNSATTNNCT 467
OY 386 ILENCRSS 395
DB 468 TILQGYQDGS 477

RESULT 4
US-10-231-778-2
Sequence 2, Application US/10231778
GENERAL INFORMATION:
APPLICANT: Billoreau, Pierre
APPLICANT: Chaudhury, Abdul M.
APPLICANT: Dennis, Elizabeth S.
APPLICANT: Koltunow, Anna M.G.
APPLICANT: Luo, Ming
APPLICANT: Peacock, William J.
TITLE OF INVENTION: Method for inducing seed development by down-regulating
FILE REFERENCE: 72-98A
CURRENT APPLICATION NUMBER: US/10/231,778
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: 09/398,237
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 60/101,184

QY 334 ADGHISMACEAFSEYKELHRYSLFWCMRLFLIKLNHGLVDSATINNCNTILLENCRN 393
DB 573 ADHVPWACEAFSRHGHGELLENLALMGMRMTIKLNHSLLSARMTDTCNRILDDIKN 632

RESULT 6

US-10-437-963-114092
Sequence 114092, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Bouharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 114092
LENGTH: 201
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1.pep
US-10-437-963-114092

Query Match 26.0%; Score 617.5; DB 6; Length 201;
Best Local Similarity 62.6%; Pred. No. 5.8e-50;
Matches 114; Conservative 26; Mismatches 39; Indels 3; Gaps 1;

QY 223 AINASSLDAGAKV--ILTSEAVVPATKTRKLSARSEARSHLLQKROFYSHRVOPTMA 279
DB 20 SVNAAVSVDPANSLHSGNSLAPVLOFGTKRKLVSERARRRROLQKROFYSHRVOPTMA 79
QY 280 LEOVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFHLNNSFVRKORYADGHIS 339
DB 80 LEOVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFHLNNSFVRKORYADGHIS 139
QY 340 MACEAFSEYKELHRYSLFWCMRLFLIKLNHGLVDSATINNCNTILLENCRN 399
DB 140 MACEAFSEYKELHRYSLFWCMRLFLIKLNHGLVDSATINNCNTILLENCRN 399
QY 400 TN 401
DB 200 KN 201

RESULT 7

US-10-424-599-283985
Sequence 283985, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 283985
LENGTH: 242
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1.pep
US-10-424-599-283985

Query Match 24.2%; Score 575.5; DB 6; Length 242;
Best Local Similarity 48.2%; Pred. No. 7e-46;
Matches 123; Conservative 38; Mismatches 63; Indels 31; Gaps 6;

QY 145 DKEFEFSLCSKPKRRRORGGRRNTTRKLCFPLDPSLTNGTITLNDGRRGLGY 204
DB 13 DRYVOTFEFCCKPKLRRTADQS---LKNA-VGLESPAGGTD--ILLEKDG----- 59
QY 205 PEATRLAGOFEMTNSIPPAIHSLDAGAVILT--SEAVVPATKTRKLSARSEARSHL 262
DB 60 -----ISATIRSPDRSDVSQMSDDOAVLOFATKTRKLSERPPRNST 104
QY 263 LLOKROFYSHRVOPTMALEQVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFHLN 322
DB 105 FLKROFYSHRVOPTMALEQVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFHLN 164
QY 323 WNSFVRKORYADGHISACEAFSEYKELHRYSLFWCMRLFLIKLNHGLVDSATIN 382
DB 165 WNSFVRKORYADGHISACEAFSEYKELHRYSLFWCMRLFLIKLNHGLVDSATIN 224
QY 363 NCNTILLENC-RNNSD 396
DB 225 DCNITILEYORNSD 239

RESULT 8

US-10-198-723-6
Sequence 6, Application US/10198723
GENERAL INFORMATION:
APPLICANT: Dang, Van-Dinh
APPLICANT: Okamoto, Jack
TITLE OF INVENTION: CHIMERIC POLYPEPTIDES MODULATING
FILE REFERENCE: 11696-043001
CURRENT APPLICATION NUMBER: US/10/198,723
CURRENT FILING DATE: 2002-07-05
PRIOR APPLICATION NUMBER: US 60/303,654
NUMBER OF SEQ ID NOS: 100
SOFTWARE: Fastseq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 692
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-198-723-6

Query Match 21.5%; Score 510.5; DB 6; Length 692;
Best Local Similarity 27.3%; Pred. No. 4.1e-39;
Matches 154; Conservative 53; Mismatches 121; Indels 363; Gaps 8;

QY 81 KTEVREDCSCPCFGLGSGFKGLQFHNSHDLFEFEFLFEYQ--TYNVSYKLNSTF 138
DB 4 KAEVVENSCPCFLIPGEGHGLQHLKSSHDKFEFYRAERDGHPEVDVSVKSDITKE 63
QY 139 E-BEGSDDDKFEFSLCSKPKR-KRORGGRRNTTRKLCFPLDPSLTNGTITL 195
DB 64 GVLKDDVGNPQLSPLTFCSSKRNORORDSDSNVKKLVLLMELDDLDPRTENDSTHV 123
QY 196 NDGN----- 199
DB 124 NDDVNSPPRAHSSEKISDILTTOLAIESSEPKVPHVNDGVNSPPRAHSAEKNEST 183
QY 200 ----- 199
DB 184 HVNDDDDVSSPPRAHSLKNESTHVNDNISPPRAHSSKNESTHVNDDEVSPPTRS 243
QY 200 -----RG----- 201
DB 244 SKETSDILTTTPAIVEPEPKVRGSRKROLAKRYKAROTPAIAESSEPKVLYHND 303
QY 202 --LQVPAETELAGFEWTSNIPPAIHSS----- 228


```

Db 304 NVSPPEAHSLEKASDILTTTQPAIAESSEPKVPHVNDVNSTPRAHSSKKNKSTRNV 363
QY 229 ----- 228
Db 364 DNVSPPKTRSSKTSIDILTTTQPTIAESSEPKVPHVNDVNSTPRAHSSKKNKSTRN 423
QY 229 -----LDAGARVILITSEAVP----- 244
Db 424 DNIPSPKTRSSKTSNLTFTQPAIAESSEPKVPHVNDVNSTPRAHSSKKNKSTHK 483
QY 245 -----ATKTKLSAESSEARSHLLQKR 267
Db 484 DNASLPKTRSSKTSIDILATTPAKAEPSEPKVTRVSRKRELHAERCEAKRLELRGR 543
QY 268 QFYHSHRYQPMALQVMSDRDSEDDVDVADFEEDROMLDFVDYNDKDEKOPMHLMSFV 327
Db 544 QFYHSHRYQPMALQVMSDRDSEDDVDVADFEEDROMLDFVDYNDKDEKOPMHLMSFV 603
QY 328 RKQRYIADGHSIMWCAEAFSRFEKELHRYSLFWCWRLEFLIKLMNHGLVDSATINNCMTI 387
Db 604 RKQRYIADGHSIMWCAEAFSRFEKELHRYSLFWCWRLEFLIKLMNHGLVDSATINNCMTI 663
QY 388 LENCNNSDITTTNNNNVDRPSDNTNNNN 418
Db 664 L-----LSNDEAGQFTSGSAANANN 684

```

RESULT 9 US-10-424-599-254263

```

; Sequence 254263, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254263
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(166)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1.pep
US-10-424-599-254263

```

Query Match 20.6%; Score 491; DB 6; Length 166;
Best Local Similarity 60.2%; Pred. No. 4e-38;
Matches 100; Conservative 21; Mismatches 39; Indels 6; Gaps 3;

```

QY 1 MCORNCRAKSSPEEISIDENLLIYCKPRLYNIFHLSISGSPSLRCLPCLANKIGAKRR 60
Db 1 MCORNSPVYHNGEETIADDESLLIYCKPVELNIIYRRLOQPSFLRCLYKTIARRRR 60
QY 61 KRSRTGVVFNKDCNNLTOKTEVEDSCSPFCMSLCSFKLQFHLNSSHDLFEFEFKL 120
Db 61 RLRA-GIYFNVRDHYNLIRKTEVEDSCSPFCMLQCSFKGLRHLNLSHDLFNFGFV 119
QY 121 FEEYTVAVSVKLSFTIEEGSD--DKFEPSLCKSPKRROR 163
Db 120 TEDYQAVNVSVKIN--IFEKENVADGVIPOSOTFFFCSPRRRRRK 163

```

RESULT 10
US-10-424-599-183487
; Sequence 183487, Application US/10424599

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 183487
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_136701C.1.pep
US-10-424-599-183487

```

Query Match 11.3%; Score 269; DB 6; Length 102;
Best Local Similarity 52.1%; Pred. No. 1.9e-17;
Matches 50; Conservative 14; Mismatches 32; Indels 0; Gaps 0;

```

QY 269 SEDVDVDDVADFEEDROMLDFVDYNDKDEKOPMHLMSFVRKQRYIADGHSIMWCAEAFSR 348
Db 7 SOQEVDDHDDVEDPATYHLHLAHSVDEKXFMHGXCHSLKRRKSVCPHGVWMAEAFXOF 66
QY 349 YEKELHRYSLFWCWRLEFLIKLMNHGLVDSATINNC 384
Db 67 YKELILSLAFWRWRLEFLIKLMNHGLHARIMNXC 102

```

RESULT 11 US-10-424-599-216298

```

; Sequence 216298, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216298
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37346C.1.pep
US-10-424-599-216298

```

Query Match 10.8%; Score 256.5; DB 6; Length 116;
Best Local Similarity 41.2%; Pred. No. 3.4e-16;
Matches 47; Conservative 26; Mismatches 38; Indels 3; Gaps 2;

```

QY 293 VDDVADFEEDROMLDFVDYNDKDEKOPMHLMSFVRKO--RVYADGHSIMWCAEAFSRFE 350
Db 1 MNEGAQELDRRL-AIIDAPVHAEFLSLMAAFIKKHGRVLADGHIMWCAEAFYTKHS 59
QY 351 KELHRYSLFWCWRLEFLIKLMNHGLVDSATINNCYTIENCNNSDITTTNNNN 404
Db 60 AFAQSNLSLAWMRMRLEFLIKLMNHGLKSTFVACNNIIEQCKRQNDPQROND 113

```

RESULT 12
US-10-437-963-199908

Sequence 199908, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 199908
LENGTH: 151
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_9542C.1.pep
US-10-437-963-199908

Query Match 10.7%; Score 255.5; DB 6; Length 151;
Best Local Similarity 42.4%; Pred. No. 6, 1e-16;
Matches 67; Conservative 15; Mismatches 39; Indels 37; Gaps 5;
DB 4 QNCRAK---SSPEEISTDENLILCKPVRLYNIFHLRSLGNPSFLPCLNYKIGAKRR 60
DB 25 QEGAKDMSESPYSVSYND-----VPSSLTHIRLS----- 58
QY 61 KRSSTGVNFYKDCNNTLOKTEVRDSCSPGSMGSGFGLOPHLNSHDLFEFEFKL 120
DB 59 -----GNVLENNYNNYNTMOKTEVEDSCPECLVPGCSFGGLGCHLASHDLHYEFWI 113
QY 121 FEERYVNSVKNLSFIFE--EGSDDDKFEPPSLCSK 156
DB 114 SEEQAVNSLKTDSWRTELAEV-DPRHQTFSYRSR 150

RESULT 13
US-10-425-114-70148
Sequence 70148, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovacic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 70148
LENGTH: 394
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-2MELB73067D07_FLI.pep
US-10-425-114-70148

Query Match 10.5%; Score 249; DB 6; Length 394;
Best Local Similarity 20.8%; Pred. No. 9, 5e-15;
Matches 70; Conservative 17; Mismatches 29; Indels 220; Gaps 2;
QY 3 RNCRAKSSPEEIVSTDENLILCKPVRLYNIFHLRSLGNPSFLPCLNYKIGAKRR-- 60
DB 55 RQLRALSPDEKTLAEKDLALYCKPVELYNIQRRAMKNLFLIORCLLVIHARRKKRI 114

QY 61 ----- 60
DB 115 QTTISLGSSTWTELQTHYVPLVYLARPTSNLSLRGSHPIYFRKCHLTSFSEHGKD 174
QY 61 -----KSRST----- 65
DB 175 NSEATFIIPDKSLSTSRACHNDIIFISCGVQSGNEDNCSGNHVEDSLQMLEGKCSW 234
QY 66 ----- 65
DB 235 GKIPNLASSLSCVNLISGHIVELASKVTMRPSFLPEKPLEDSCITFCSHKVDVGS 294
QY 66 -----GMVFENYKDCNNTLOKT 82
DB 295 YKLOCMGSAQEGARDMSLSPYSYNDVPSSLSDIIRLSGNVLEFVKYNTMTQET 354
QY 83 EYREDSCSPGSMGSGFGLOPHLNSHDLFEFEFKL 118
DB 355 EYTEDSCSPGFCYVACGSGFGLOCHLNSHDLFEFEFK 390

RESULT 14
US-10-437-963-184120
Sequence 184120, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 184120
LENGTH: 263
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(263)
OTHER INFORMATION: unsure at all Xaa locations
OTHER INFORMATION:
OTHER INFORMATION: Clone ID: PAT_MRT4530_81143C.1.pep
US-10-437-963-184120

Query Match 8.0%; Score 190; DB 6; Length 263;
Best Local Similarity 53.8%; Pred. No. 2e-09; 18; Indels 0; Gaps 0;
Matches 35; Conservative 12; Mismatches 12

QY 1 MCRCRAKSSPEEIVSTDENLILCKPVRLYNIFHLRSLGNPSFLPCLNYKIGAKRR 60
DB 3 MCRHOPRARLSPDEQLAERESFALYCKPVELYNIQRRAMKNLFLIORCLLVIHARRKK 62
QY 61 KRSST 65
DB 63 RSLIT 67

RESULT 15
US-10-424-599-250979
Sequence 250979, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 250979
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(195)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MPT3847_68663C.1.pap
US-10-424-599-250979

```

```

Query Match      6.7%; Score 159; DB 6; Length 195;
Best Local Similarity 56.9%; Pred. No. 1.1e-06;
Matches 29; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 11 SPEVISTDENTLIYCKPVRLYNIFHLRSLGNSFLPRLCLANKKIGAKRRK 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 22 SAEELAAESLSIYCKPVELYNILQRRMRNPSPFLQRCIHYRIKAKRRK 72

```

Search completed: June 13, 2003, 15:37:12
Job time : 56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 19, 2003, 22:10:23 ; Search time 3586 Seconds

(without alignments)
3611.475 Million cell updates/sec

File: US-09-890-220-2

Sequence: 1 MCRQNCRAKSSPEVISTDE.....INNNNVNDKNNSDKVIK 445

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: . 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV-xlh
-Q/cgn2.1/USPTO-spool/US09890220/runat_13062003_144851_16477/app.query.fasta.1.583
-DB=GenEmbl -QFWT=fastap -SUFFIX=trge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human4.0.cdi -LIST=45
-DOCCALIGN=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220 &CGN_1.1.2456 &runat_13062003_144851_16477 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2378	100.0	1722	6	AX032817	AX032817 Sequence
2	2378	100.0	1722	6	AX032888	AX032888 Sequence
3	2370	99.7	1700	8	AF284500	AF284500 Arabidops
4	2363	99.4	1722	6	AX032934	AX032934 Sequence
5	2343.5	98.5	1722	8	AF284501	AF284501 Arabidops
6	2285	96.1	1354	8	AY063047	AY063047 Arabidops
7	2285	96.1	1715	6	AX032818	AX032818 Sequence
8	2285	96.1	1715	6	AX032891	AX032891 Sequence
9	2285	96.1	1721	6	AY034902	AY034902 Arabidops
10	2258.5	95.0	1737	6	AX032894	AX032894 Sequence
11	1813.5	76.3	6338	6	AX032890	AX032890 Sequence
12	1798.5	75.6	91660	8	ATRPSP5LE2	AF180942 Arabidops
13	1659	69.8	5895	6	AX032893	AX032893 Sequence
14	1659	69.8	199548	8	ATCHRIV45	AL161545 Arabidops
15	1659	69.8	201471	8	ATRCAT7	297342 Arabidops
16	923	38.8	2279	8	AB053171	AB053171 Arabidops
17	589	24.8	2284	8	AF096096	AF096096 Arabidops
18	451.5	19.0	558	8	AX032900	AX032900 Sequence
19	400.5	16.8	8508	8	AB053262	AB053262 Arabidops
20	400.5	16.8	87180	8	AB023044	AB023044 Arabidops
21	368.5	15.5	325	6	AX032898	AX032898 Sequence
22	311	13.1	463	6	AX032896	AX032896 Sequence
23	306.5	12.9	135495	2	OSJN01007	AL607099 Oryza sat
24	265	11.1	107436	8	AC006068	AC006068 Arabidops
25	264.5	11.1	6081	8	AF096095	AF096095 Arabidops
26	250	10.5	2040	8	AB053265	AB053265 Arabidops
27	250	10.5	199548	8	ATCHRIV45	AL161545 Arabidops
28	250	10.5	201471	8	ATRCAT7	297342 Arabidops
29	242	10.2	127130	8	OSJN00040	AL606603 Oryza sat
30	240	10.1	113355	8	OSJN00100	AL606654 Oryza sat
31	222	9.3	3346	9	AK074333	AK074333 Homo sapi
32	221	9.3	3942	9	BC015704	BC015704 Homo sapi
33	221	9.3	4441	6	AX032902	AX032902 Sequence
34	221	9.3	4441	6	AX337789	AX337789 Sequence
35	221	9.3	4441	9	D63881	D63881 Human mRNA
36	219	9.2	3648	3	AF149047	AF149047 Drosophill
37	219	9.2	4055	3	AY069809	AY069809 Drosophill
38	208	8.7	83538	2	AC018338	AC018338 Drosophill
39	208	8.7	167816	3	AC009381	AC009381 Drosophill
40	208	8.7	249805	3	AE003515	AE003515 Drosophill
41	178.5	7.5	57873	2	AC010021	AC010021 Drosophill
42	175	7.4	56152	2	AC116963	AC116963 Dictyoste
43	168	7.1	5769	3	AF104350	AF104350 Dictyoste
44	161	6.8	171187	2	AC116960	AC116960 Dictyoste
45	155.5	6.5	12464	2	AC115604	AC115604 Dictyoste

RESULT 1

ALIGNMENTS

AX032817
 LOCUS AX032817 1722 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 57 from Patent WO0046358.
 ACCESSION AX032817
 VERSION AX032817 GI:10279793
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM
 REFERENCE
 AUTHORS Dean, C., West, J. and Johanson, U.
 TITLE Plant gene
 JOURNAL Patent: WO 0046358-A 57 10-AUG-2000;
 DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
 JOHANSON URBAN (SE)

FEATURES
 source 1..1722
 location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="landsberg erecta VRN2 cDNA"

BASE COUNT 538 a 339 c 348 g 497 t
 ORIGIN

Alignment Scores:
 Pred No.: 1,46e-209 Length: 1722
 Score: 2378.00 Matches: 443
 Percent Similarity: 100.008 Conservative: 0
 Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 100.008 Indels: 0
 Gaps: 6

US-09-890-220-2 (1-445) x AX032817 (1-1722)

QY 1 MetcysarglnasnancysargalalyserserprogluvaliliserthrAspGlu 20
 Db 231 ATGTGTAGGAGCAATGTCGCGCAATCTCCACGAGAGTATTCACACGATGAG 290
 QY 21 AsnleuauiletyrcyslyspovalargleutyAsnillepHehisleuAArgSerleu 40
 Db 291 AATCTCTTGATATATGTAAACCTGTCCACATAATACATCTTCACCTCGCTCTTA 350
 QY 41 GlyAsnProserPheleuproarqCysleuAsnTYRylleGlyAlaIysArgLysArg 60
 Db 351 GGCACCCATCGTTCTCCACAGATCTTGAACTCAAAATGGAGCAAGCGCAAGA 410
 QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTYRylAspCysAsnAsnThrleuGln 80
 Db 411 AAGTCAGATCTACTGGAGATGGTAGTTTCAACTATTAAGATTTGTAATACACTACAG 470
 QY 81 LysThrGluValAlaArgGluAspCysSerCysPropheCysSerMetleuCysGlySerPhe 100
 Db 471 AAAACGTAACTTAGGAGAGATGTTCTTCTTCATTTGCTCATATCTGTGTAGCTTC 530
 QY 101 LysGlyLeuGlnPheHisleuAsnSerSerHisAspLeuPheGluPheGluPheLysleu 120
 Db 531 AAGGCGCTGCAATTTCTATTTGATTCATTCATGATTTATTTGATTTGAGTTCAAGCTT 590
 QY 121 PheGluGluTYRylGlnThrValAsnValSerValLysleuAsnSerPheIlePheGluGlu 140
 Db 591 TTCGAGAAATACACAGACAGTAAATGTTCTGTAACCTTAATTCCTTATATTTGAGGAA 650
 QY 141 GluGlySerAspAspAspLysPheGluProPheSerleuCysSerLysProArgLysArg 160
 Db 651 GAAGGAGATGATACATTAATTTAGCCCTTCTCTCTGCTCGAAACCTCGTAAGCGG 710
 QY 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLysValCysPheleuProleuAsp 180
 Db 711 AGACAAAGAGGTGCGCAATACACACAGACACTTAAAGTATGCTTTTACCTTGAT 770
 QY 181 SerProSerleuThrAsnGlyThrGluAsnGlyIleThrleuLeuAsnAspGlyAsnArg 200
 Db 771 TCACCCAGTTTAACTTAATGACACAGAAATGAAATCACCTACTTAATGATGAAACGGT 830

QY 201 GileuGlyTYRylProGluAlaThrGluLeuAlaGlyInPheGluMetThrSerAsnIle 220
 Db 831 GGTTAGATATATCCGAGAGCAAGACCTGCTGGCAAAATTTAGATGACCGCAACATTT 890
 QY 221 ProProAlaIleAlaHisSerSerleuAspAlaGlyAlaLysValIleleuthrSerGlu 240
 Db 891 CCACAGCCATPAGCCCACTTCTCTGTGACGCTGTCTTAATATTTATGACAGCGAA 950
 QY 241 AlaValAlaProAlaThrLysThrArgLysleuSerLalaLysSerGluAlaArgSer 260
 Db 951 GCTGTGTCCTGCTGCTACTTAACAGAAAGATTAATCTGACGATCAGAGCTAGAGCTA 1010
 QY 261 HisleuLeuGlnLysArgGlnPheTYRylHisSerHisArgValGlnProMetAlaLeu 280
 Db 1011 CACCTACTTCTTCAGAAACGCCAATTCATCATTTCTACAGACTCCAGCCATGCGCTT 1070
 QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
 Db 1071 GAGCAAGTAATGTCTGACCGGAGATAGCGAGATGAATGCATACATGTTCCAGATTTT 1130
 QY 301 GluAspArgGlnMetleuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 Db 1131 HisleuTYRylAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr 340
 QY 321 HisleuTYRylAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr 360
 Db 1191 CATCTTTGCACTGCTTTGTAAAGAAACAAAGGCTTATACAGATGATGATATCTTGGG 1250
 QY 341 AlaCysGluAlaPheSerArgPheTYRylLysGluLeuHisArgTYRylSerSerleuPhe 360
 Db 1251 GCATGTGAAGCATTTTCAAGATTTTACAGAAAGATGACCGGTACTCATCATCTCTTC 1310
 QY 361 TTPCysTYRylArgLeuPheleuIleLysleuTYRylAsnHisGlyLeuValAspSerAlaThr 380
 Db 1311 TCGTGTGGAGATTTGTTTGTATTAACATGAGAACCATGACCTTGTCCACGACCCAC 1370
 QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
 Db 1371 ATCAACAACCTGCAATACCATCATCTCGAAATTTGCCGTATAGCTCAACACCCACAC 1430
 QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 Db 1431 AACCAACAACAAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490
 QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
 Db 1491 GATCATCCCATGACATTAACACAAACACATGTTGACACAGACACATTAACAGCACA 1550
 QY 441 AspLysValIleLys 445
 Db 1551 GACCAAGTAAATTA 1565

RESULT 2
 LOCUS AX032888 1722 bp DNA linear PAT 22-SEP-2000
 DEFINITION Sequence 1 from Patent WO0044918.
 ACCESSION AX032888
 VERSION AX032888 GI:10279825
 KEYWORDS
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS Dean, C. and Gendall, A.
 TITLE Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 JOURNAL Patent: WO 0044918-A 1 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)
 FEATURES
 source 1..1722
 location/Qualifiers

BASE COUNT 538 a 339 c 348 g 497 t
 ORIGIN /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"

Alignment Scores:
 Pred. No.: 1,46e-209 Length: 1722
 Score: 2378.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-890-220-2 (1-445) x AX032888 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 Db 231 ATGTGTAGGCAAAATTTGTCGCGCAAAATCTCCACGGAGGAAGTGATTTCAACTGATGAG 290
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 Db 291 AATCTCTGATATATTGTTAACTGCTGACATATATACATCTTTCCACTTGCCTCTCTA 350
 QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 Db 351 GGCAACCCATCGTTCTTCTCCAGATGCTTGAACTACAAAATTGGAGCAAGCGCAAAAGA 410
 QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 Db 411 AAGTAAAGATCTACAGGGATGGTGTAGTTTCAACTAAAGCATTTGTAATACACATTCACG 470
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
 Db 471 AAAACTGAAGTTAGGAGAGATGTTCTTCTCCATTTTGCCTCATGCTATGCTAGTCTTC 530
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluLysLeu 120
 Db 531 AAGGGGCTGCATTTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590
 QY 121 PheGluGluTyrGluThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 Db 591 TTGCAAGATACAGACAGTAAATGTTTGTAAATCTTAATCTTCAATTTTCCATTTTGGAGA 650
 QY 141 GluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 Db 651 GAAGAGATGATGAGATTAATTTGAGCCCTCTCTCTGCTCGAAACCTGTAAGCGG 710
 QY 161 ArgGlnArgGlyGlyAlaArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
 Db 711 AGACAAAGAGGTGGCAGAAATTAACACAGAGACACTTAAAGTATGCTTTTACCGTGCAT 770
 QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
 Db 771 TCACCCAGTTTAACATATGACACAGAAATGAATCACTCACTAATATATGAAACCGCT 830
 QY 201 GlyLeuGlyTyrProGluValThrGluLeuAlaGlyGluPheGluMetThrSerAsnIle 220
 Db 831 GGTTTAGATATCCGAGGCAACAGAGCTTGTGTCGACATTTGAGATGACGCAACATTT 890
 QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 Db 891 CCACCGCCATAGCCCACTCTCTGACGCTGGTGCTAAAGTTATATATGACAAACGAA 950
 QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 Db 951 GGTGGGTCCCTGCTACTAAGACAAAGAAATGATATCTGACGACGATACAGGCTAAAC 1010
 QY 261 HisLeuLeuGluGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 Db 1011 CACCTACTCTTCAGAAAGCCCAATTTCTATCTTCACAGAGTCCAGCCAAATGGCGCTT 1070
 QY 281 GlnGluValMetSerAspArgAspSerGluAspGluValAlaAspAspValAlaAspPhe 300

Db 1071 GAGCAAGTAATGTCTGACCCGGATAGCAGATGAAGTGCATGACATGTCAGATTTT 1130
 QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAspLysAspLysGlnPheMet 320
 Db 1131 GAAGATCGCCACATGCTTGATGACTTTTGAGATGTGAATGAAGATGAAGCAATTCAGT 1190
 QY 321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerThr 340
 Db 1191 CATCTTTGAACTCGTTTGTAAAGAAACAAAGGTTATGACATGCTATGCTCTTG 1250
 QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGlyLeuHisArgTyrSerSerLeuPhe 360
 Db 1251 GCATGGAAGCAATTTTCAAGATTTTAAGAGAAAGAGTTTCCACCGTTACTCATCAGCTTTC 1310
 QY 361 TrpCysTrpArgLeuPheLeuLeuLysLeuThrAsnHisGlyLeuValAspSerAlaThr 380
 Db 1311 TGGTGTGGAGATGTGTTTGAATTAACATATGAAACCAAGCAAGCACTGTCACACCCAC 1370
 QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
 Db 1371 ATCAACAACATGCATATGCATCTCCGAAATTCGCCGTAATAGCTCAGACACACACAC 1430
 QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 Db 1431 AACAAACAACAAGATGGATGCTGCCAGTGACTCAACACCAACAACAATTAACATTTGTG 1490
 QY 421 AspHisProAsnAspLysAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
 Db 1491 GATCATCCCAATGACATTAACAACAAGCAAAATGTTGACAAACAGCAATTAACACACA 1550
 QY 441 AspLysValIleLys 445
 Db 1551 GACAAAGTAATTTAA 1565

RESULT 3
 AF284500 1700 bp mRNA linear PLN 21-NOV-2001
 LOCUS Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete cds.
 DEFINITION
 ACCESSION AF284500
 VERSION AF284500.1 GI:16945787
 KEYWORDS
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 1700)
 Gendall, A.R., Levy, Y., Y., Wilson, A. and Dean, C.
 The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis
 Cell 107 (4), 525-535 (2001)
 JOURNAL
 MEDLINE 21575875
 PUBMED 11719192
 REFERENCE 2 (bases 1 to 1700)
 Gendall, A.R., Levy, Y., Y. and Dean, C.
 Direct Submission
 JOURNAL Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
 FEATURES
 source
 1..1700
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="IV"
 /map="adjacent to RPP5"
 /note="ecotype: Landsberg erecta"
 1..1700
 /gene="VRN2"
 227..1564
 /gene="VRN2"
 /codon_start=1
 /product="vernalization 2 protein"
 /protein_id="AAU32135.1"

```

/db xref="GI:16945788"
/translation="MCRONCRASSPSPVETISIDENLLIYCKPVRILYNTFHRISGNS
ELPRCLNLIKIGARERKRSRSTGMYVENVYDCNNITLOKTEVEDDSCPPCSMLCSFG
LOPHINSNDLPPEPKISEPYOTVAVSKLNSFIIEEGSDDDFEFFSLCSKPRK
ROGGRNNRRLKVCFLPDSFSLNGSLNGLITLNDGNKGLIPEATELGGPEKMS
NTPPALHNSLSDGAKVILITSEAVPATKRLSLKSEANSHLLORQFHSRVO
PALQVMSRSDSEDDVDVADFEEDRDLDDVDNDEKQFMHMSFVKORVIA
DGHISMACNFSRFEKELHRTSSLPCWCRFLIKLNMHGLVDASATLNNCTVLENCR
NSSDTTNNNNNSVDRPSDNTNNNNIVDPNDINNNNNVNDKONNSRDKYIK"
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="a"
1195
variation
/ gene="VRN2"
/ note="VRN2-1 mutation results in TGG Trp to TGA stop"
/ replace="a"
1399
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="g"
1487
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="a"
1488
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="g"
BASE COUNT      533 a      334 c      347 g      486 t
ORIGIN
Alignment Scores:
Pred. NO.:      7.83e-209      Length:      1700
Score:          2370.00      Matches:      444
Percent Similarity: 99.78%      Conservative: 1
Best Local Similarity: 99.78%      Mismatches: 0
Query Match:    99.66%      Indels:      0
DB:             8      Gaps:        0
US-09-890-220-2 (1-445) x AF284500 (1-1700)
QY      1 MetCysArgGlnAsnGlySerProGluGluValIleSerThrAspGlu 20
Db      227 ATGTGTAGGAGAGATTGCGCGCAATCCTCACCGAGAAAGATTCAACTGATGAG 286
QY      21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
Db      287 AATCTCTGATATATATGTAACCTGTCGACTATATACATCTTCACTTCGCTCTCTA 346
QY      41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
Db      347 GGCACCCATCGTTTCTTCCAGATGCTGGAACATCAAAATTGGAGCAAAACGCAAAAGA 406
QY      61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
Db      407 AACTCAAGATCTACTGGATGTGTAATTTCAACTTAAGCATTTAATACACATTTCAG 466
QY      81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
Db      467 AAAACTGAAGTTRAGGAGATGTTCTGTCTCTCATTTCTCTATAGCATGAGGAGGTTTC 526
QY      101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
Db      527 AAGGGGCTCAATTTCAATTTGAATTCATCTCATGATTTTGAATTTGAATTTCAACTT 586
QY      121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
Db      587 TCGAAGAAATACCAACAGATTAATGTTTCTGTAATAAACTTAATTCCTCATATTTGAGAA 646

```

```

QY      141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
Db      647 GAAGGAAGGAGAGACGATAAATTGAGCCCTCTCTCTGCTGCAAACTGTAAGCCG 706
QY      161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
Db      707 AACACAAAGGCGGACGAGAAATTAACACAGGAGACTTAAATATGCTTTTACCGTGGAT 766
QY      181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
Db      767 TCACCCAGATTAACTAAATGACGACGAAATGAGATCCCTACTATATATGTAAGAACCGT 826
QY      201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
Db      827 GGTTAGGATATCCGAGGACAAACAGACTGCTGACATTTGATGATGACAGCAACAT 886
QY      221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
Db      887 CCACAGGCATAGGCCACTCTTCTCTGAGCGTGTGCTAAATTTATTTGACCAAGCGAA 946
QY      241 AlaValAlaProAlaThrLysThrArgLysLeuSerLagLysSerGluAlaArgSer 260
Db      947 GCTGTGTCCTCCGCTACTTACACACAGAAAGTATCTCTGAGCATCAGAGCTAGGCTAGAAC 1006
QY      261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
Db      1007 CACCTACTTCTTCAAGAAACGCCATTTCTATCATTTCTACAGAGTCCAGCCATTCGCGCTT 1066
QY      281 GluGlnValMetSerAspAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
Db      1067 GAGCAAGTAAATGTCTGACCGGATAGCAGAGATGTAAGTATGATGATGATGATGATGATGAT 1126
QY      301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
Db      1127 GAAAGATCCGACAGATGCTTGATGACTTGTGATGTGAATTAACATGAAAGCAATTCATG 1186
QY      321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspLysHisIleSerTrp 340
Db      1187 CATCTTTTGAACTCGTTTGTAAGAAACAAAGGGTATATGACAGTGTATATTTCTGCGG 1246
QY      341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
Db      1247 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGACTTGCACCGCTACTCATCATCTCTC 1306
QY      361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrAsnHisGlyLeuValAspSerAlaThr 380
Db      1307 TGGTGTGAGATTTGTTTGATTAATCTATGACACATGACATGACTTGTGACTCAGCCACC 1366
QY      381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrTrp 400
Db      1367 ATCAACAACTGCATATTCATCTCGTGAATATGCGCTGAATATGCTCAACACCAACCCACC 1426
QY      401 AsnAsnAsnAsnSerValAspArgProSerAspSerAspThrAsnAsnAsnIleVal 420
Db      1427 AACAAACAACAACATGTGAGATCGGCCAGTACTCAACCAACAACAATAATTAATTTGTG 1486
QY      421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
Db      1487 GTTCATCCCAATGACATTAACAAACAAACAAATGTTGACAAACAGACATTAACAGCAGA 1546
QY      441 AspLysValIleLys 445
Db      1547 GACCAAGTAATTAATA 1561

```

```

RESULT 4
AX032934
LOCUS      AX032934      1722 bp      DNA      linear      PAT 21-SEP-2000
DEFINITION Sequence 47 from Patent WO0044918.
ACCESSION  AX032934
VERSION     AX032934.1 GI:10279842
KEYWORDS    thale cress.
SOURCE

```


ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1722)
AUTHORS Dean, C. and Gendall, A.
TITLE Methods and means for modification of plant characteristics using the vernalization gene vrn2
JOURNAL Patent: WO 0044918-A 47 03-Aug-2000;
DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD (GB)

FEATURES
source Location/Qualifiers
1..1722
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 539 a 339 c 347 g 497 t
ORIGIN

Alignment Scores:
Pred. No.: 3.52e-208 Length: 1722
Score: 2363.00 Matches: 444
Percent Similarity: 99.788 Conservative: 0
Best Local Similarity: 99.788 Mismatches: 1
Query Match: 99.37% Indels: 0
DB: Gaps: 0

US-09-890-220-2 (1-445) x AX032934 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerProGluGluValIleSerThrAspGlu 20
DB 231 AGTGTAGGCAAGATTGTCGCGGAAATCTCACCCGAGAGATGATTTCACTGTATGAG 290
QY 21 AsnIleuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerIleu 40
DB 291 AATCTCTATATATTGTAACCTGTTCCACTATATTAACATCTTACCTCGCTCTCTA 350
QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 351 GGCACCCATCGTTCTTCCAGATGCTTGAACCTACAAATGAGCAAGCCCAAGCA 410
QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 411 AAGTCAAGATCTACTGGATGGATGTTTCACTAATAGGATTTGTAACACATTTACG 470
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 471 AAAACTGAGATTAGGAGGATTTGTTCTTGTCCATTTTGTCTATAGTATGTGTACTTC 530
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 531 AAGGGGCTGCATTTTCATTTGAATCATCATGATTTATTTGAATTTGAGTTCAAGCTT 590
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGln 140
DB 591 TTCGAAAGATFACACAGACACTTAATGTTCTGTAATAAATTAATCTTCAATATTGAGGA 650
QY 141 GlnGlySerAspAspLysPheGluPhePheSerLeuCysSerLysProArgLysArg 160
DB 651 GAAGGAGATGATGACATTAATTTTGAAGCCCTTCTCTGCTGCGAAGCCCTGTAAGCGG 710
QY 161 ArgGlnArgGlyValArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 711 AGCAAAAGAGGTGGCGAATAAACAACAGAGACTTAAGTATGCTTTTACCGTTGAGT 770
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
DB 771 TCACCCAGTTTAACTATGACGACAGAAATGGAATCACCCTAATGATGGAACCGCT 830
QY 201 GlyLeuGlyTyrProGluAlaThrGluAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 831 GGTATTAGATATCCCAAGGCAACAGAGCTTGCTGGCAATTTTGAGATGACCAACACATT 890
QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGln 240

DB 891 CCACGAGCATAGCCACACTCTCTGACGGCTGGGTAAAGTTATATGACAAAGCGGA 950
QY 241 AAlaValProAlaThrIleThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 951 GGTGTGTCTCCCTGCTACTAAGACAGAAAGTTATCTGCGACGAGCAAGGCTTAACAC 1010
QY 261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1011 CACCTACTCTTCAGAAACGCCAATTTCTATCTTCTCACAGAGTCCAGCAATGGCGCTT 1070
QY 281 GlnGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
DB 1071 GAGCAAGTATAGTCTGACCGGGATGACGAGATGAAGTGCATGATGTCAGATTTT 1130
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB 1131 GAAGATCGCCAGATGCTTATGATGATCTTGGATGTGATGAATGAATGAATGAATGAAT 1190
QY 321 HisLeuTyrPheSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB 1191 CATCTTTGAACTCGTTTGTAAAGAAACAAAGGTTATAGCCAGATGCTATATCTTGG 1250
QY 341 AlaCysGluAlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPhe 360
DB 1251 GCATGGAAGCATTTTCAAGATTTTACGAGAAAGTGCACCGTTACTCATCATCTCTTC 1310
QY 361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrPheHisGlyLeuValAspSerAlaThr 380
DB 1311 TGGTGTGGAGATGTTTTGTATTAACATATGAAACAGAGCTTGTCCAGCAGCCACC 1370
QY 381 IleAsnAsnCysAsnThrIleLeuGlnAsnCysArgAspSerSerAspThrThrThr 400
DB 1371 ATCAACAATGCAATATCATCTCCGAGATTCGCCGTAATAGCTCACACACACACACAC 1430
QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerSerAsnThrAsnAsnAsnIleVal 420
DB 1431 AACCAACAACAAGTGTGATGTCCTCCACTGACTCAAAACCAACAACATTAACATTTG 1490
QY 421 AspHisProAspAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB 1491 GATCATCCCAATGACATTAACAACAAGAACAAATGTTGACAAAGGACATTAACACGACA 1550
QY 441 AspLysValIleLys 445
DB 1551 GACAAAGTATTTAA 1565

RESULT 5
AF284501 1722 bp mRNA linear PLN 21-NOV-2001
LOCUS Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
DEFINITION cds, alternatively spliced.
ACCESSION AF284501.1 GI:16945789
VERSION AF284501
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
REFERENCE 1 (bases 1 to 1722)
AUTHORS Gendall, A.R., Levy, Y.Y., Wilson, A. and Dean, C.
TITLE The VERNALIZATION 2 gene mediates the Epigenetic Regulation of Vernalization in Arabidopsis
JOURNAL Cell 107 (4), 525-535 (2001)
MEDLINE 21575875
PUBMED 11719192
REFERENCE 2 (bases 1 to 1722)
AUTHORS Gendall, A.R., Levy, Y.Y. and Dean, C.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
FEATURES Location/Qualifiers

TITLE
JOURNAL

COMMENT

Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., Shin, P., Southwick, A., Shinozaki, K., Davis, R. W., Ecker, J. R. and Theologis, A.
Direct Submission
Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Banb, J., Banno, F., Chang, E., Dale, J. M., Goldsmith, A. D., Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Toriumi, M., Wu, H. C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M. C., Miranda, M., Nguyen, M., Palm, C. J., Shin, P., Southwick, A., Davis, R. W., Ecker, J. R. and Theologis, A.
Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.
Location/Qualifiers

FEATURES
SOURCE

1..1354
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="4"
/clone="F09871"
/note="This clone is in PUNI 51.
ecotype: Columbia"
1..1354
/gene="At4g16845"
1..1323
/gene="At4g16845"
/codon_start=1
/evidence="experimental
/product="unknown protein"
/protein_id="AAL34221.1"
/db_xref="GI:17104665"

CDS
gene

1324..1354
/gene="At4g16845"
BASE COUNT 431 a 264 c 287 g 372 t
ORIGIN
3'UTR
NTSVTNNNNNSVDDHSDSNNTNNNNIVDHPNDIKNNVNDKNNRDX"

Alignment Scores:

Pred. No.: 3.94e-201 Length: 1354
Score: 2285.00 Matches: 428
Percent Similarity: 97.51% Conservative: 3
Best Local Similarity: 96.83% Mismatches: 9
Query Match: 96.09% Indels: 2
DB: 8 Gaps: 1

US-09-890-220-2 (1-445) x AY063047 (1-1354)

QY 1 MetCysArgGlnAsnGlyAlaGlySerSerProGluGluValIleSerThrAspGlu 20
DB 1 ATGTGTAGGCAAAATTTGTCGGCGGAATCTTCACCGAGGAGTATTTCAACTATGAG 60
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40

DB 61 AATCTCTGATATATGTAACCTGCTGACATATTAACATCTTCACCTCGCTCTTA 120
QY 41 GlyAsnProSerPheLeuProArgCysLeuSerTyrIleGlyAlaLysArgLysArg 60
DB 121 GGCAACCATGCTGTTCTGCCAAGATGCTTGAACCTCAAAATTTGGGCAAGCGCAAGA 180
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnThrLeuGln 80
DB 161 AAGTAAAGATCTACTGGATGGTATGTTTCAACTTAAGATGTAATTAATACATTACA 240
QY 81 LysThrGluValArgLysAlaPysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 241 AGAAGTGAAGTATAGGAGAGATGTTCTTCTGTCATTTCTCATATTCAGTTCGCTTC 300
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheLysLeu 120
DB 301 AAGGGCGTCGCAATTTCAATTTGAATTCATCTCATATTTGAATTTGAGTTCAGCTT 360
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 361 TTGGAAGATACAGACAGATATATGTTCTGTAACCTTAATTCCTTCATATTTAGGA 420
QY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 421 GAAGAGATGATGATGATTAATTTGAGCCTTCTCTGCTGCTGCAAACTGTAAGCT 480
QY 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 481 AGACAAGAGGCGGAGAAATACACAGGAACTTAAGTATGCTTTTACCGTGGAT 540
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
DB 541 TCACCCAGTTTACGTAATGAGCACAGAAATGAAATGTCCTGCTCAATGATGAAACCT 600
QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 601 GGTGTTAGATATCCGAGGCAACAGAGCTGCTGCAATTTGAGATGACTGCAACAT 660
QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 661 CCACAGGCATAGCCACCTCTCTGAGCGTGTGTAAGTATATTAACAACCA 720
QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 721 GCTGTGTCCTCTGCTACTAAGACAGAAAGTTATGCTGACCGATGACAGGCTGAAGC 780
QY 261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 781 CACCTACTCTTCAGAAACGCCAATTCATATCTTCACAGAGTCCAGCAATGGCGCT 840
QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
DB 841 GAGCAAGTAATCTCGATCGAGTAGCGGAGTGAAGTCGAGAGAGATGTTCCACATTTT 900
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB 901 GAAAGATCGCCATGCTTGATGACTTTGATGATGATTAAGATGAAGAAAGCAATTCATG 960
QY 321 HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr 340
DB 961 CATCTTTGGAACCTGTTGTAAGAAACAAAGGTTATAGCAGATGTCATATCTTGG 1020
QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
DB 1021 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTATCTACATCTTTC 1080
QY 361 TrpCysTyrArgLeuPheLeuIleLysLeuTyrPasnHisGlyLeuValAspSerAlaThr 380
DB 1081 TGGTGTGAGATGTTTGTGATTAACATAGAACCATGACCTGTGACTCAGCCACC 1140
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400

Db 1141 ATCAACAACGTAATCCATCCCTCGAGAAATTCGCTAATACCTCA-----GTCACTAAC 1194
 QY 401 AAsAsAsAsAsSerValAsPaRgProSerAspSerAsnThrAsnAsnAsnVal 420
 Db 1195 AACACACACACAGCTGTGATCCCTCCAGCTCAACACACACACAAATACATTTG 1254
 QY 421 AsPHisProAsnAsp1LeAsnAsnValAsPaSAsnValAsPaSAsnSerArg 440
 Db 1255 GATCATCCGATGACATTAATAAACAAGAACAAATGTTGACAAACAGACAAATACACGAGA 1314
 QY 441 AsPlys 442
 Db 1315 GACCAAG 1320

RESULT 7
 AX032818
 LOCUS AX032818 1715 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 58 from Patent WO0046358.
 AX032818
 VERSION AX032818.1 GI:10279794
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM
 REFERENCE
 AUTHORS Dean,C., West,J. and Johanson,U.
 TITLE Plant gene
 JOURNAL Patent: WO 0046358-A 58.10-AUG-2000;
 DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
 JOHANSON URBAN (SE)

FEATURES
 source
 1. 1715
 location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Columbia VRN2 cDNA"
 BASE COUNT 536 a 326 c 353 g 500 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5,42e-201 Length: 1715
 Score: 2285.00 Matches: 428
 Percent Similarity: 97.518 Conservative: 3
 Best Local Similarity: 96.838 Mismatches: 9
 Query Match: 96.098 Indels: 2
 DB: Gaps: 1

US-09-890-220-2 (1-445) x AX032818 (1-1715)

QY 1 MetCyArGlnAsnCysArGAlaLysSerSerProGluGluVal11SerThrAspGlu 20
 Db 233 ATGTGTAGGCGAGAAATGTCGCGCAATCTCACCGAGAAAGATTTCAACGATGAG 292
 QY 21 AsnLeuLeu1LeTyCysLysProValArgLeuTyRsn1LePheH1sLeuArgSerLeu 40
 Db 293 AATCTGTGATATATTGTAACCTGTGACTATATAACAATCTTCACTTCGCTCTCTA 352
 QY 41 G1YAsnProSerPheLeuProArgCysLeuAsnTyRys1LeG1AlaLysArgLysArg 60
 Db 353 GGAACCCATCTCTTCTCGCAAGAGCTTGAACACAAATTTGGGCAAGAGCGAAGAGA 412
 QY 61 LysSerArgSerThrG1LysMetValAlaPheAsnTyRysAspCysAsnAsnThrLeuGln 80
 Db 413 AATGTCAAGATCTACTGGGTGGTGTACTTTCAACATATAAGGATTTGTAATATACATTACA 472
 QY 81 LysThrGluValAlaGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
 Db 473 AGAACTGAAGTAGGAGAGATTTGTTCTTCATTTTGCCTATGCTATGTGGTAGCTTC 532
 QY 101 LysGlyLeuGlnPheH1sLeuAsnSerSerH1sAspLeuPheGluPheGluPheLysLeu 120
 Db 533 AAGGGGCTGCATTTCAATTTCAATTTCAATTCATCATCATGATTTATTTGAATTTGAGTTCAAGCTT 592
 QY 121 PheGluGluTyRGlThrValAsnValSerValLysLeuAsnSerPhe1LePheGluGlu 140

Db 593 TTGGAAAGAAATACCAAGAGTTAAATGTTGTGAATTAATCTTCAATTTTGAAGAA 652
 QY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 Db 653 GAAGGAATGATGATGATTAATTTTGGACCCCTTCTCTCTGCTCGAAACCTGTAAGCGT 712
 QY 161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
 Db 713 AGCAAAAGAGGTGGCAGAAATACACCGAGAGACTTAAAGTATGCTTTTACCGTTGAT 772
 QY 181 SerProSerLeuThrAsnGlyThrGluAsnGly1LeThrLeuLeuAsnAsp1LysAsnArg 200
 Db 773 TCACCCAGTTTAGCTAATGGCAGAAAATGGAATGGCCCTGCTGAATGATGAAACCGT 832
 QY 201 GlyLeuGlyTyRProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsn1Le 220
 Db 833 GGTTAGGATATCCCGAGGACACAGAGCTTGTGACAAATTTGAGATGACTAGCAACATT 892
 QY 221 PropAla1LeAlaH1sSerSerLeuAspAlaGlyAlaLysVal1LeuThrSerGlu 240
 Db 893 CCACAGCCATAGCCCACTCTTCTGTGACGCTGGTGCTAAAGTTATATTACACCGCA 952
 QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 Db 953 GCTGTGCTCCCTGCTACTAAGACAGAAAGTTATCTGCTGACGATCAGAGCTTGAAGC 1012
 QY 261 HisLeuLeuLeuGlnLysArgGlnPheTyRHisSerHisArgValGluProMetAlaLeu 280
 Db 1013 CACCTACTTCTTCAAGAAAGCCAAATCTATCATCTTCAAGAGTCCAGCAATGGCGCT 1072
 QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
 Db 1073 GAGCAAGTAATGTGTGATCGGATAGCGAGATGAAAGTGATGATGATGATGATGATGAT 1132
 QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 Db 1133 GAAGATCCCGCAGATGCTGTGATGACTTGTGATGATGATGATGATGATGATGATGATGAT 1192
 QY 321 HisLeuThrAsnSerPheValArgLysGlnArgVal1LeAlaAspGlyH1s1LeSerTrp 340
 Db 1193 CATCTTTGGAACTCGTTGTGAAAGAAACAAAGGCTTATGACAGATGGCTATTCCTTGG 1252
 QY 341 AlaCysGluAlaPheSerArgPheTyRGlLysGluLeuH1sArgTyRSerSerLeuPhe 360
 Db 1253 GCATGTGAAGATTTTCAAGATTTTGAAGAAAGAGTTGCACTGTACATCAACACTCTTC 1312
 QY 361 TrpCysTrpArgLeuPheLeu1LeLysLeuTrpAsnH1sGlyLeuValAspSerAlaThr 380
 Db 1313 TGGTGTGGAGATTTGTTTGAATTAACATATGAAACCATGAGACTTGTGACTCACGCCAC 1372
 QY 381 IleAsnAsnCysAsnThr1LeuGluAsnCysArgAsnSerSerAspPheThrThrThr 400
 Db 1373 ATCAACAACGCAATACCATCTCGAAGATGGCGTAATACCTCA-----GTCACTAAC 1426
 QY 401 AsnAsnAsnAsnSerValAsPaRgProSerAspSerAsnThrAsnAsnAsnVal 420
 Db 1427 AACACACACACAGCTGTGGATCATCCAGAGATCTAAACACCAACAAATACATTTGTG 1486
 QY 421 AsPHisProAsnAsp1LeAsnAsnValAsPaSAsnValAsPaSAsnLysAsnAsnSerArg 440
 Db 1487 GATCATCCGATGACATTAATAAACAAGAACAAATGTTGACACMACAGACAAATACACGAGA 1546
 QY 441 AsPlys 442
 Db 1547 GACCAAG 1552

RESULT 8
 AX032891
 LOCUS AX032891 1715 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 4 from Patent WO0044918.
 AX032891
 ACCESSION AX032891.1 GI:10279827

KEYWORDS

SOURCE

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 1715).

Dean, C. and Gendall, A.

METHODS AND MEANS FOR MODIFICATION OF PLANT CHARACTERISTICS USING

THE VERNALIZATION GENE VRN2

Patent: WO 0044918-A 03-AUG-2000;

DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD

(GB)

FEATURES

source

1..1715

/db_xref="taxon:3702"

/db_xref="taxon:3702"

BASE COUNT

536 a 326 c 353 g 500 t

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Best Query Match:

DB:

US-09-890-220-2 (1-445) x AX032891 (1-1715)

1 MetcysArgGlnAsnGlySerProGluValIleSerThrAspGlu 20

223 ATGTGTAGGCAAGATTTGCGCGCAATCTCAGCGAGAGATTTCAACTATGAG 292

21 AsnLeuLeuIleThrCysLysProValArgLeuThrAspGluValIleSerThr 40

223 AATCTCTGATATATTTGAAACCTCTGACATATTAACATCTTTCACCTTCTCTCA 352

41 GlyAsnProSerPheLeuProArgSerLysLeuValIleGlyValLysArgLysArg 60

333 GCGAACCCATGCTTCTGCGCAAGATCTTGAACAAATTTGGGCAAGCGCAAGA 412

61 LysSerArgSerThrGlyMetValValPheAsnThrLysAspCysAsnAsnThrLeuGln 80

413 AAGTCAAGATCTACCTGGAGATGTTTCAACTTAAGATTTGAATATATCATACCA 472

81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100

473 AGAAGTGAAGTTAGGAGAGATTTGCTTCTGCTATTTGCTATATGTTAGCTTC 532

101 LysGluLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120

533 AAGGGCTGCAATTTCAATTTGATTCATCTCATGATTTATTTGAAATTTGAGTTCAAGCTT 592

121 PheGluGluThrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140

593 TTGGAAGATACCAACAGATTAATGTTCTGTAATAAATTAATTCCTTCATATTAGGAA 652

141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160

653 GAAAGAAAGTGAATGAATTAATTTAGCCCTTCTCTCTGCTCGAACAACCTTGTAAGCGT 712

161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180

713 AGACAAAGAGGTGGAGAAATTAACACCGAGAGACTTAAGATTAATGCTTTTACCGTTGAT 772

181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200

773 TCACCCAGTTTAAGTAATGCGACAGAAATATGATTTGCGCTGATGATGATGAAACCGT 832

201 GlyLeuGlyThrProGluValThrGluLeuValGlyGlnPheGluMetThrSerAsnIle 220

833 GCTTTAGATATATCCCGAGCAACAGAGCTTGTGACAAATTTGAATGATGATGCAACATTT 892

221 ProProIleIleAlaHisSerSerLeuAspLysArgValIleLeuThrSerGlu 240

893 CCACGAGCCATAGCCACCTCTTCTTGAGACGCTGCTGCTAAAGATTAATTAACAACCGAA 952

241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluLysArgSer 260

953 GCTGTGTCCTCGCTACTAAGCAAGAAAGTATTTGCTGACGATGACAGGCTTGACAGC 1012

261 HisLeuLeuLeuGlnLysArgGlnPheThrHisSerHisArgValGlnProMetAlaLeu 280

1013 CACCTACTTCTTCAAGAAAGCCATTTATCTATCTCTCAGAGATGACAGGCTGACGCTT 1072

281 GluGlnValMetSerAspArgSerGluAspGluValAspAspValAlaAspPhe 300

1073 GAGCAAGTAATGCTGATCGGATGCGAGATGAGATGAGATGAGATGAGATGAGATGAG 1132

301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320

1133 GAAAGATCGCCAGATCTTGATGATCTTGTGATGATGATGATGATGATGATGATGATG 1192

321 HisLeuThrPheSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340

1193 CATCTTTGAGACTCGTTTGTAGAAAGCAAGAGTTATAGCAGATGCTATCTCTTGG 1252

341 AlaCysGluAlaPheSerArgPheThrGluLysGluLeuHisArgThrSerSerLeuPhe 360

1253 GCATGTGAAGTATTTTCAAGATTTTACGAAAGAGTTGACCTGTTACTCATCATCTTTC 1312

361 TrpCysTrpArgLeuPheLeuIleLysLeuThrPheAsnHisGlyLeuValAspSerAlaThr 380

1313 TGTGTGTGAGATGTTTGTGATTAACATGAGAACCATGAGCTTGTGACATCAGCACACC 1372

381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400

1373 ATCAACAACATGCAATATACATCTCGAGATTTCCGTAATATCACTCA-----GTCACTAAC 1426

401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420

1427 AACCAACACACAGGTGATGATATCCAGTGCCTCAACACCAACACCAATATATGTTG 1486

421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440

1487 GATCATCCGAAATGACATTAATAAACAAGAAACATGTTGACCAACAGACATTAACAGCAGA 1546

441 AspLys 442

1547 GACAAAG 1552

RESULT 9

AY034902

LOCUS

AY034902

DEFINITION

Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete

cDNA.

VERSION

AY034902.1

KEYWORDS

FLI CDNA.

SOURCE

Arabidopsis thaliana.

ORGANISM

Arabidopsis thaliana.

REFERENCE

1 (bases 1 to 1721)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Ban, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,

Carnilini, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,

Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K.,

Davis, R.W., Becker, J.R. and Theologis, A.

Arabidopsis Full Length cDNA Clones

REFERENCE

2 (bases 1 to 1721)

TITLE

JOURNAL

Unpublished

REFERENCE

AUTHORS

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Kosemura, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinozaki, A., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologos, A.

TITLE

Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN Arabidopsis Full Length cDNA": Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Kosemura, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shino, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologos, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologos, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES

source

1. 1721

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="4"

/clone="RAF105-08-K19 (R09811)"

/note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SerI insert."

ecotype: Columbia

1. 1721

/gene="At4g16845"

1. 231

/gene="At4g16845"

232. 1534

/gene="At4g16845"

/codon_start=1

/evidence="experimental"

/product="unknown protein"

/protein_id="AAK59409.1"

/db_xref="GI:14334422"

/translation="MCRQNCRAKSSPEEVISTDENLTYCKPYRLXNIETRLSAGNS

FLPRLNLTAKRKRSRTGMYVFNKDKNNLTQRTREDDSCFCSMLGSPKRG

LOPHLNSHDLFEFFKLEBYOTVNSVKLNSIFEEDSDDKPEPFLSCGPKR

RQGRNNTRRLKVCFLDEPDLANGENGLANGINGLGYEPETELACGFEKTS

NIPLAIASSLDAGAKVLTTEAVAPATKRLTSAESESRSLLLOKROFYGEHVA

PALEQVSDRSDSEDDVDVAFEDROMLDEVDYDKDKRQPMHLNNSVFRQRTYA

DGHISMACEVFSREFYKELHCYSLLFCWRLFLIKLNHGLVYSATINNNITLENCR

NTSVNNNNNSVSDHPSDSTNNNNIYVHPNDIKKNNVYDKNNNSRDK"

1555. 1721

/gene="At4g16845"

3'UTR

BASE COUNT

548 a 324 c 353 g 496 t

ORIGIN

Alignment Scores:

Pred. No.: 5 45e-201

Score: 2285.00

Percent Similarity: 97.51%

Best Local Similarity: 96.83%

Query Match: 96.09%

DB: 8

Indels: 2

Gaps: 1

US-09-890-220-2 (1-445) x AY034902 (1-1721)

1 MetCysatggtglnasgcysatgalaayserseerprogluivaallieserthraspglu 20

|||||

232 ATGTGTAGGACAGATTGTCGCCGCAAAATCTCACCCGAGACAGTATTCACACTGAG 291
 21 AsnLeuLeuLeuTyrCysLysProValArgLeuTyrAsnLeuPheHisLeuArgSerLeu 40
 292 AATCTTGAATATATATGAAACCTGTCGATATATACATTCATACCTGCTCTCA 351
 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 352 GCAACCCATCGTCTTCTCCCAAGATGCTTGAACATCAAAATTTGGGCAAAAGCCAAAGA 411
 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 412 AAGTCAAGATCTACTGCGATGATGTTTCACTAATAGATAGATATATATACATCA 471
 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 472 AGACTGAAGTTAGGAGAGATGTTCTTCTCATTTTGTCTATGCTATGTGTAGCTTC 531
 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluLysLeu 120
 532 AAGGGCTGCATTTTCATTTGATATTCATCATGATATTTATTTGATTTGATTTCAAGCTT 591
 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 592 TTGGAAGATACCAAGACAGTTAATGTTCTGTAAACTTATTCCTCATATTTAGAGAA 651
 141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
 652 GAAGAAATGATATGATATATATTTGACCCCTCTCTCTGCTGCAAACTGTAAAGCT 711
 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
 712 AGACAAAGATGTCAGAAATTAACACAGAGACTTAAGTATGCTTTTACCGTTGAT 771
 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
 772 TCACCAATTTAGCTAATATGACAGAAATGAAATGAAATGCTGCTGATATGAAACCGT 831
 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
 832 GGTATAGGATATCCCGAGGACAGACAGCTTGTGAGCAATTTGAGATGCTAGCAACATT 891
 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 892 CCACAGGACATAGCCCACTCTCTCTGAGCGTGGCTGAAGTATATATTAACAACCGAA 951
 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 952 GCTGGGCTCCCTGCTACTAGACAAAGAAATTTATCTGTGACGATCAAGGCTGAAAGC 1011
 261 HisLeuLeuGlnLysArgLysPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 1012 CACTACTCTTTCGAAAGCCCAATTCATCATCTCAACAAGTCCAGCAATGCGCTT 1071
 281 GluIleValMetSerAspArgAspSerGluAsnGluValAspAspValAlaAspPhe 300
 1072 GAGCAAGATATGCTGATTCGCGATACGAGATGAAGTGAAGTGAAGTGTGCAATTTT 1131
 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 1132 GAAGATGCGCAGATGCTTGTGATGACTTTGTGAGATGAATTAAGAAAGCAATTCATG 1191
 321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerThr 340
 1192 CATCTTTGAACATCGTTTGAACAAACAAAGGTTTAAGCAGATGTCATATCTCTGG 1251
 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
 1252 GCATGTAAAGTATTTTCAAGATTTTACGAAAGAGTTCGCTGTTACTCATCTCTTC 1311
 361 TrpCysTrpArgLeuPheLeuLysLeuThrAsnHisGlyLeuValAspSerAlaThr 380
 1312 TGGTGTGAGATTTGTTTGAATTAACATATGAAACATGACCTTGTGACTGACCAACC 1371

QY 381 ileaasnCyAsnThrileuGlusncysArasSerSeraSphthrrthr 400
 DB 1372 ATCAACACATGCAATACCATCTCCAGAAATTCCTTAATCTCA-----GTCACTAAC 1425
 QY 401 AsnaAsnaAsnSerValAsparProSerAspSerAsnThrAsnaAsnaSnileVal 420
 DB 1426 AACACACACACAGTGTGATCATCCAGTACTCAACACACACACAAATTAATGTTG 1485
 QY 421 AsphSPROASnAspIleasAsnlySAsnaSnValAspAsnlySAspAsnaSnSerArg 440
 DB 1486 GATCATCCGAATGACATATAAAACAAACAAATGTTGACACACAGACAAATTAACAGACA 1545
 QY 441 AspLys 442
 DB 1546 GACAAAG 1551
 RESULT 10
 AX032894 1737 bp DNA linear PAT 21-SEP-2000
 LOCUS Sequence 7 from Patent WO0044918.
 DEFINITION AX032894
 ACCESSION AX032894
 VERSION AX032894.1 GI:10279829
 KEYWORDS
 SOURCE
 ORGANISM
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1737)
 Dean, C. and Gendall, A.
 Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 Patent: WO 0044918 A 7 03-AUG-2000;
 DEAN CAROLINE (GB); GENDALL ANTHONY (GB); PLANT BIOSCIENCE LTD
 (GB)
 FEATURES
 source 1..1737
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 543 a 330 c 359 g 505 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,536-198 Length: 1737
 Score: 2258.50 Matches: 428
 Percent Similarity: 95.78% Conservative: 3
 Best Local Similarity: 95.11% Mismatches: 9
 Query Match: 94.97% Indels: 10
 DB: 6 Gaps: 2
 US-09-890-220-2 (1-445) x AX032894 (1-1737)
 QY 1 MetCysArgGluAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 233 ATGTAGGACGAATTTGTCGCGCAATCTCACCGAGAGATGATTCAACTGATGAG 292
 QY 21 AsnIleuIleIleTyCysLysProValArgLeuIleuIleuIleuIleuIleuIleuIleu 40
 DB 293 AATCTCTGATATATGTAAACCTGTCACATATTAACATCTTTCACCTTGCTCTCTCA 352
 QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnIleuIleuIleuIleuIleuIleuIleu 60
 DB 353 GGCACCCATCTCTTCTGCGCAAGATGCTTGAACCTACAAATTTGGGCAAGCGCAAAACA 412
 QY 61 LysSerArgSerThrGlyMetValValPheAsnIleuIleuIleuIleuIleuIleuIleu 80
 DB 413 AAGTCAAGATCTACAGGATGATGATTCAACTAAGATTTGTAATATACATTACA 472
 QY 81 LysThrGluValArgGluAspCysSerCysPropheCysSerMetLeuCysGlySerPhe 100
 DB 473 AGAACTGAAGTTAGGAGATGTTCTGTCATTTGCTATGCTATGATGATGCTTC 532

QY 101 Lys-----GlyLeuGlnPheHisLeuAsnSerHisAspLe 113
 DB 533 AAGCGGCGCACTATTACACACTGAGCGGCTGCAATTTCTATTGATTCATCTCATGATTT 592
 QY 113 nPheGluPheGluPheLysLeuPheGluGluIleuIleuIleuIleuIleuIleuIleuIleu 133
 DB 593 ATTGAATTTGAGTTCAAGCTTTTGGAGAAATACAGACAGTAAATGTTTCTGTAAACT 652
 QY 133 uAsnSerPheIlePheGluGluGluGlySerAspAspLysPheGluIleuIleuIleuIleu 153
 DB 653 TAATTCCTTCATATTGAGAGGAAGGAGATGATATGATTAATTTGAGCCCTTCCTCTCT 712
 QY 153 uCysSerLysProArgLysArgArgLysArgLysArgLysArgLysArgLysArgLysArg 173
 DB 713 CTGCTCGAAACCTCGTAAACCGTACGACAAAGAGGTGCGCAAAATTAACACGAGACTTAA 772
 QY 173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyIleuIleuIleuIleuIleu 193
 DB 773 AGTATGCTTTTACCGTTGAGATTACCCAGTTTACCTAATGCGACAGAAATGGAATTGC 832
 QY 193 rIleuLeuAsnAspGlyAsnArgGlyLeuGlyIleuIleuIleuIleuIleuIleuIleuIleu 213
 DB 833 CTGCTGAAATGATGAAACCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 892
 QY 213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyVal 233
 DB 893 ATTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
 QY 233 aLysValIleIleuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAl 253
 DB 953 TAAAGTTATATTAACAAACCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
 QY 253 aGluArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyHisSerHis 273
 DB 1013 TGAAGCATCAGAGGTAGAGGACCATCTTCTTGAAGAACCCCAATTTTATCATCTCA 1072
 QY 273 sarGValGlnProMetAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluVal 293
 DB 1073 CAGAGTCCAGCCCAATGCGCTTGACCAAGTATATGCTGATGCGGATAGAGAGATGAACT 1132
 QY 293 LAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
 DB 1133 CGATGACGATGTTGCGAGATTTTGAAGATCGCCAGATGCTGTGATGATGCTGATGATG 1192
 QY 313 nLysAspGluLysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArgValAl 333
 DB 1193 TAAAGATGAAAGCAATTCATGCTATGCTGTTGAGTGTGAAGAAACAAAGGGTAT 1252
 QY 333 eAlaAspGlyHisIleSerThrPalAcysGluAlaPheSerArgPheTyGlyLysGlyLe 353
 DB 1253 AGCAGATGCTCATCTCTTGCGCATGTCGAATATTTTCAGATTTTTCAGAAAGAGATT 1312
 QY 353 uHisArgTySerSerLeuPheThrPcysTrpArgLeuPheLeuIleuLysLeuTrpAsnHis 373
 DB 1313 GCACGTGATCATGATCATCTCTGCTGTTGAGATGTTTGTGATTAACATATGAAACCA 1372
 QY 373 sGlyLeuValAlaAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAs 393
 DB 1373 TGGACTTGGCGACTAGCCATCAACAACTGATATACATCTCTGGAATTTGGCGGTA 1432
 QY 393 nSerSerAspThrThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAs 413
 DB 1433 TACCTCA-----GTCACTAAACAAACAAACAGTGTGATCCACAGTACATCAAA 1486
 QY 413 nThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAs 433
 DB 1487 CACCAACACATATACATGATGATCATCCGATGACATTAATAAACAAGAAACAAAGTGTGA 1546
 QY 433 pAsnLysAspAsnAsnSerArgAspLys 442
 DB 1547 CAACAAGCAATTAACAGCAGACAAAG 1574
 RESULT 11

OY 202 uclytyrproglualathrglualeuaglylnpnehiuemetthrserasnlepropr 222
 |||||
 DB 4021 AGGATATCCCGAGGACAGACAGACTGCTGACAAATTTGGATGACCAGCAATTCACC 4080
 OY 222 oalailaialhisserSerleuaspalaglyalysvalilleuethrSerglualaya 242
 DB 4081 ACCCATATGCCACTCTCTCTGAGCGCTGCTAAAGTTATATTGACAAAGCGAAGCTGT 4140
 OY 242 lvalproalathrlysthrArglyleuSerlialuagserglualiarq----- 259
 |||||
 DB 4141 GGTCCCTGCTACTAAGACAAAGAAAGTTATCTGTCAGCATCAGAGGCTAGAAGGTTTGT 4200
 OY 259 ----- 259
 DB 4201 TCATCATGACACCCGTCATCATTAATTAACATTCCGTGTTCACAATGTTCTCTATT 4260
 OY 260 -----Serhisleu 263
 DB 4261 ATGATTAAGTGTATATAGTACTGCATATTAAACGAGAAATTTCTTCAGCCACTACT 4320
 OY 263 uleuolnlysarglInpethrHisSerhisArgval----- 275
 |||||
 DB 4321 TCTTCAGAAACGCCAATTCATTCATCTTCACAGAGTCAGGTATCCAAAGTTCTTCACC 4380
 OY 275 ----- 275
 DB 4381 TACTCTTAGCATTTCTTTAAATTCCTCATGATGATATCTTATCAAGCATACTGTGT 4440
 OY 275 ----- 275
 DB 4441 TTGTTCTACCAAAATTTGTATTTGATCTGATGATCAACGCAAAATAGTTATGTCCA 4500
 OY 276 -----Gln-Prometa 279
 |||||
 DB 4501 TGTGTCTCCGTTTATTGCGCACTAACCAAAAAATGCATGTTTCTGTGCAAGCCATG 4560
 OY 279 lalengluglnvalmetSeraspArgspSerglunaspgluvalaspaparpvalalaa 299
 |||||
 DB 4561 GCCTTGAGCAATTAATGTGTGACCGGAGTAGCGAGATGAATGCATGACATGTTGCGAG 4620
 OY 299 spPheglunaspArg----- 303
 |||||
 DB 4621 ATTTTGAAGATCGCAGATTCATGATTTCTTCTGCGTCATTAAATAGACAACAGA 4680
 OY 304 -----Glnmetleu 306
 |||||
 DB 4681 AAATGATATGATGATTAACCTGCTAATGCGCTTTGAAACTTAAAAAGCTGCAGATGCTT 4740
 OY 307 ASPaspPhevalaspvalasnlyspglulysglInPheMetHisleuTripsnsSerPhe 326
 |||||
 DB 4741 GATGACTTTGTGATGTGAATTAAGATGAAAAAGCAATTCATGCAATCTTTGGAACTCGTTT 4800
 OY 327 valArglysglnArgval----- 332
 |||||
 DB 4801 GTAAGAAAAAAGGA-ACACTCTCTTACACATGAACAGACACAAAAAGACCTTANG 4859
 OY 332 ----- 332
 DB 4860 TCTTACATTCATACCTGCTAAATGATTTGCTTATGGAACCTTGACCTCAATATGAT 4919
 OY 333 -----llealaspgllyhisIleSerTriPalacysglualapheSer 346
 |||||
 DB 4920 TGTGATGTTTCAGGGTTATACACAGTGTCAATATTTCTTGCGCATGTGAAGCATTTTCA 4979
 OY 347 ArgpethrGlulysgluleuHisArglyrSerSerleuPheTriP----- 361
 |||||
 DB 4980 AGATTTTACGAGAAAGAGTTGACCGCTTACTCATCTCTCTCGTAATATATAGTACACC 5039
 OY 361 ----- 361
 DB 5040 AAACATATACAGACACATACATACATCAATCTTGTTCGTTTCTGAAAAAATA 5099

OY 362 -----CysTrpArgleupheleuilelysleuTriPasnHisGlyleuValasp 377
 |||||
 DB 5100 AAAATTCAGAGTGTGGAGATGTTTGTGATTAACATATGAGAACCATGATTCGAC 5159
 OY 378 SerAlathrIleasnAsnCysasnThrIleleugluAsnCysArgasnSerSerAspThr 397
 |||||
 DB 5160 TCAGCCACCATTCACAACTGCATATCCATCCCTCGAAGATTCGCTGAATAGCTCACACACC 5219
 OY 398 ThrThrThrAsnAsnAsnSerValaspArgProSerAspSerAsnThrAsnAsn 417
 |||||
 DB 5220 ACCACCACCAACACAAACAGTGTGATGCTCCAGTCAAAACCAACAAAT 5279
 OY 418 AsnIleValaspHisProAsnAspIleasnAsnLysAsnAsnValaspAsnLysAspAsn 437
 |||||
 DB 5280 AACATTGTGRTCAATCCCATATGACATTAACAAAGAAACAAATGTTGACACAGACAT 5339
 OY 438 AsnSerArgaspLysValIleLys 445
 |||||
 DB 5340 AACAGCAGACAAAGTATTA 5363
 RESULT 12
 ATRP5LE2/c 91660 bp DNA linear PLN 22-DEC-2000
 LOCUS Arabidopsis thaliana RPP5 disease resistance locus, main contig.
 DEFINITION AF180942 U97106
 ACCESSION AF180942.1 GI:6449044
 VERSION
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM
 Arabidopsis thaliana.
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 76020 to 83095)
 2 of 2
 Arabidopsis thaliana.
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 76020 to 83095)
 2 (bases 1 to 91660)
 Noel, L., Moores, T., van der Biezen, E.A., Parniske, M.,
 Daniels, M.J., Parker, J.E., and Jones, J.D.
 Pronounced intraspecific haplotype divergence at the RPP5 complex
 disease resistance locus of Arabidopsis
 Plant Cell 11 (11), 2099-2112 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 9212464
 2 (bases 1 to 91660)
 Noel, L., Moores, T., van der Biezen, E.A., Parniske, M.,
 Daniels, M.J., Parker, J.E., and Jones, J.D.
 Pronounced intraspecific haplotype divergence at the RPP5 complex
 disease resistance locus of Arabidopsis
 Plant Cell 11 (11), 2099-2112 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 10559437
 3 (bases 76020 to 83095)
 Coleman, M.J., Parker, J.E., Szabo, V., van der Biezen, E.A.,
 Daniels, M.J., and Jones, J.D.G.
 Direct Submision
 Submitted (10-APR-1997) Sainsbury Laboratory, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, UK
 4 (bases 1 to 91660)
 Noel, L., Moores, T., van der Biezen, E.A., Parniske, M., Daniels, M.J.,
 Parker, J.E., and Jones, J.D.G.
 Direct Submision
 Submitted (26-AUG-1999) Sainsbury Laboratory, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, United Kingdom
 On Nov 18, 1999 this sequence version replaced gi:109274.
 COMMENT
 FEATURES
 source
 1. 91660
 /organism="Arabidopsis thaliana"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="C18a"
 5354..9673
 /gene="La-1"
 /pseudo
 join(5354..5829,5977..7057,7241..7510,7610..8071,

CDS

Best Local Similarity: 43.28% Mismatches: 6
 Query Match: 75.63% Indels: 570
 DB: 8 Gaps: 13

US-09-890-220-2 (1-445) x ATPP5LE2 (1-91660)

QY 1 MetcysArgInaSnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 88263 ATGTCTAGGCAAAATGTGCGCGCAAAATCCTCAGCGGAGGATATTCAACTATGAG 88204
 QY 21 AsnLeuLeuIleTyrCysLysProValAlaArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 88203 AATCTTGATATATTGTAACCTGTGCACTATATTAACATCTTCACCTCGCTCTCA 88144
 QY 41 GlyAsp----- 42
 DB 88143 GGCAACGATATGATTTGCCCTTCTCTCATCATATATTAAGCTCAGTAATCTTCATCTCTG 88084
 QY 42 ----- 42
 DB 88083 TGTATATCACCCACTAATATAGTTTGATTTGCTAAGCTGATATATGCTGATTCATGGCGA 88024
 QY 43 -----ProSerPhe 45
 DB 88023 GTGTGTGCTTCTTTTGTCTCTTAATTTGAACCTGTGTGTGTGTGCGAGCCATGCTT 87964
 QY 46 LeuProArgCysLeuAsnTyrIleGlyAlaLysArgLysSer----- 62
 DB 87963 CTTCAGATGCTTAACACTACAAATTTGAGCAAGCGCAAGAAAGTA-TGTTTCTT 87905
 QY 62 ----- 62
 DB 87904 CTGAATGAGCTGCTACAGTATATGTTATTATCTTACTCTTAATATGAGAGCTGAG 87845
 QY 62 ----- 62
 DB 87844 ACCATTTTATCTTTGTTGATGATATGACATATGATGTTTCTTCTTGTTCATG 87785
 QY 63 -----ArgSerThrGlyMetVal 68
 DB 87784 CTATTAACCTTACATTTTATATAATGTTGTTGGTTAGGTCAAGATCTAGGAGTGA 87725
 QY 69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 DB 87724 GTTTCACTAATAAGATTTGTATATACACATTTACAGAAACTGAAAGTTAGCTTTTCT 87665
 QY 83 ----- 83
 DB 87664 GTTCGTGCAAAATTCGATGTCATGCTATGTTTCTTCTAGATATTTGTTATTACTA 87605
 QY 84 -----ValArgGlnAspCysSerCysProPheCysSerMetIle 96
 DB 87604 TTTTTCCTGATTTGTATGATGAGTGAAGGAGATTTGTTCTTCATTTTGTCTATGCT 87545
 QY 96 uCysGlySerPheLys----- 101
 DB 87544 ATGTGTAGCTTCAAGGTGGGCAACTATTACAACAGAGTTTCTTCCGGGGCCTTTTCATA 87485
 QY 101 ----- 101
 DB 87484 TCTAACACTGTGAATGTCTACTGCTGTTTCATGCTATATCTTCACTGTTGGTTACAT 87425
 QY 101 ----- 101
 DB 87424 ATTTTGTGTTTGTGTTTGTCTTCTCTACCTTTTTCGAACCTGAGTGTGCTTATCT 87365
 QY 102 -----GlyLeuG 104
 DB 87364 GAGAAACATGTCCAGATGAGGCTTACACCAATTTCTGTGCTATAGCAGGGGCTGTC 87305
 QY 104 InPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu----- 120
 DB 87304 AATTCATTTGAAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGGATATGTGTTTGA 87245

QY 120 ----- 120
 DB 87244 TGGAAATCTTGATTTGCTATGCCCTTTATTAATGAGTTATGTTAAAAAGGCTCTTC 87185
 QY 121 -----Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
 DB 87184 CTATGTAGCTTTGCGAAGAAATACAGACAGTAAATGTTTGTAAACTTAATTCCTTC 87125
 QY 137 IlePheGluGlu----- 140
 DB 87124 AATATTGAAGT-CAGTACTTAACTTGTTAAATGGGAAATCCGATAGCTGGTGAATA 87066
 QY 141 -----GluGlySerAspAspLysPhe 148
 DB 87065 TTTTGTATATTCATCTTATTTGTACTAGAGAAAGAGATGATGAGATTAATTT 87006
 QY 149 GluProPheSerLeu----- 153
 DB 87005 GAGCCCTTCTCTCTGTTAGTAACTTCAGAAACCTTCGATTAATACCTTAATAGCAGTA 86946
 QY 154 -----CysSer 155
 DB 86945 ACTCCTGCTTCTTCTGTCAGTACATCTCTGTAATTCACCAACCATATGTTTTCAGCTCG 86886
 QY 156 LysProArgLysArgArgGlnArgGlyLysArgAsnThrArgArgLeuLysValCys 175
 DB 86885 AATCTCGTAACCGGAGACAGAGAGGTGCGACAAATATACACAGGAGCTTAAAGTATGC 86826
 QY 176 PheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeu 195
 DB 86825 TTTTACCGTTGATTTCAACCCAGTTTAACATATGACAGAAATGGAATCACCTACTT 86766
 QY 196 AsnAsp----- 197
 DB 86765 AATGATGTGAATATCATATCTTCTCTGCGTCTGCTGCTTGAACCTTCAATATAC 86706
 QY 197 ----- 197
 DB 86705 ACAAGACATACAAATGCCCTGATTTGTTAGTTTGTACTTCTCTCGCATCTCTTTC 86646
 QY 198 -----GlyAsnArgGly 201
 DB 86645 GAGGATATGTATACAGAACTGATGACAAATTAATGACATGCTACAGAAACCGTGTG 86586
 QY 202 LeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIlePro 221
 DB 86585 TTAGATATACCCGAGGCAACAGAACTGCTGACAAATTTGAGATGACCGCAACATTTCCA 86526
 QY 222 ProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAla 241
 DB 86525 CCAGCCATAGCCCACTCTTCTTGACGCTGCTGTATTAAGTATATGCAAGCGAAGCT 86466
 QY 242 ValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259
 DB 86465 GTGTGCCCTGCTACTAAGACAGAAAGTATCTGTGACCGATTCAGAGGCTGAAGGTTT 86406
 QY 259 ----- 259
 DB 86405 GTTCATCATGACCCCGTCATCATTAATTAACATTCCTGTTGTACAAATGTTCTCTCA 86346
 QY 260 -----SerHisLeu 262
 DB 86345 TTATGATATAGCTTTATATAGTACTGCAATTAATACCGAAGAAATTTCTTCAGCCACCTA 86286
 QY 263 LeuLeuGlnLysArgGlnPheTyrHisSerHisArgVal----- 275
 DB 86285 CTTCCTCAAGAACGCCAATTTATCATCTCTACAGAGTCCAGGTGATCCAGTCTCTTCA 86226
 QY 275 ----- 275
 DB 86225 CCTACTTCTAGCAATTTTCTTTAAATGCTCATGATGATATCTTATCAAAACATACCTTG 86166

```

OY 275 ----- 275
Db 86165 GTTGTTCATCCAAATTTGTAATTTGATGATGATCAACGCAAAATAGTATGTC 86106
OY 276 ----- Gln-Prom 278
Db 86105 CATGTGTCTCCGTTTATTTGACCAACCAAAATGACGTTTCTGTGCAAGCAAT 86046
OY 278 talaLeugluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAl 298
Db 86045 GGGCTTGACCAAGTATGTCGACCGGATGACGAGATGAAGTCGATGATGTC 85986
OY 298 aaspPheGluAspArg ----- 303
Db 85985 AGATTTTGAAGATGCCAGGATATCCATGATTTCTTGGCTTATTAATGACAA 85926
OY 304 ----- GlnMetL 306
Db 85925 GAAATGCTATGATGATTAACCTTGCTAATGCTTTGAACTTAAAGCTGCAATGC 85866
OY 306 euAspAspPheValAspValAsnLysAspGluLysGluPheMetHisLeuTrpAsnSerP 326
Db 85865 TTGATGACTTGTGATGTGATTAAGATGAAGATCAATTCATGATCTTTGAACTGCT 85806
OY 326 heValArgLysGlnArgVal ----- 332
Db 85805 TTGTAAAGAAAACAAAGTA-ACACTTCTCTTACATGACGACGACAAAGAACCTTA 85747
OY 332 ----- 332
Db 85746 TGTCTTACATTCATACCTCTCTAAATGATTTTGGCTTATGCACTTGACTCAATTATG 85687
OY 333 ----- IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheS 346
Db 85686 ATTGTTGATGTTTCAAGGTTTATGACGATGTCATATTTTCGGGCACTGTAACATTTT 85627
OY 346 eArgPheTrpGluLysGluLeuHisArgTyrSerSerLeuPheTrp ----- 361
Db 85626 CAAGATTTTACGAGAAAGAGTTGACACCTTACTCATCTCTCTGTAATTAAGTACA 85567
OY 361 ----- 361
Db 85566 CCAACATATACAGACATTAACATCACTATCAATCTGTTCTGTAAGAAAAA 85507
OY 362 ----- CysTrpArgLeuPheLeuLysLeuTrpAsnHisGlyLeuVal 377
Db 85506 TAAAAATTTCCAGGTGTTGACATGTTGTTTATTAACATTAAGAACATGATGATGTCG 85447
OY 377 sPseAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspT 397
Db 85446 ACTCAGCCACACATCACTCACTCAATACATCTCGAGATTCGCGTAATGCTCACACA 85387
OY 397 hrThrThrTrpAsnAsnAsnSerValAspArgProSerAspSerTrpAsnAsn 417
Db 85386 CCACACACACACACACACACACAGTGGATGCTCCAGTACCAACCAACCAACA 85327
OY 417 snAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAsp 437
Db 85326 ATAACATTTGAGTCTATCCCATGACATTAACACAAAGATGTTGACACAAAGACA 85267
OY 437 snAsnSerArgAspLysValIleLys 445
Db 85266 ATAACACGACAGACAAAGTAATTA 85241

```

```

RESULT 13
LOCUS AX032893 5895 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 6 from Patent WO0044918.
ACCESSION AX032893
VERSION AX032893.1 GI:10279828
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

FEATURES
    source
    BASE COUNT 1742 a 1027 c 1070 g 2056 t
    ORIGIN
        Alignment Scores:
        Pred. No.: 1,44e-142 Length: 5895
        Score: 1659.00 Matches: 424
        Percent Similarity: 42.25% Conservative: 12
        Best Local Similarity: 41.86% Mismatches: 577
        Query Match: 69.76% Indels: 14
        DB: Gaps: 14
        US-09-890-220-2 (1-445) x AX032893 (1-5895)
    OY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
    Db 1905 ATGTGTAGGCAAGATATGTGCGGCAAAATCTCCACCGAGAGAGATTTCAACTATAG 1964
    OY 21 AsnLeuLeuIleTrpCysLysProValArgLeuTrpAsnIlePheHisLeuArgSerLeu 40
    Db 1965 AATCTCTGATATATGTGAATGAACCTGTCGACTATTAACATCTTCCCTGCTCTTA 2024
    OY 41 GlyAsn ----- 42
    Db 2025 GCCAA-CGTATGATTTGGCTTCTCTCTCATCATTTTATAGCTAGTAATCTTTCATCTCC 2083
    OY 42 ----- 42
    Db 2084 TGTGTAGATCACCCACTAATAGTTGAGTTTGCTAAGCTGATATGCTGACTCATG 2143
    OY 43 ----- Pro 43
    Db 2144 GAGTGTGCTTCTTTGCTCCCTAATGTTTGAACCTGTTGTTGTTGCGACCA 2203
    OY 44 SerPheLeuProArgCysLeuAsnTrpLysIleGlyAlaLysArgLysSerSer 62
    Db 2204 TCGTTTCTGCCAAGATCTTGAACTACAAATATGGGGCAAAAGGCAAAAGAAAGTA-TGC 2262
    OY 62 ----- 62
    Db 2263 GTTTCCTTCTGAATGATGTTGCCACAGTATGATGATTTATTTACTTACTCTTAATATGGA 2322
    OY 62 ----- 62
    Db 2323 GCTGTGACATATTATTTACTTTGTTGAGTACATATGACATATGAATGATGTTCTTCTTTC 2382
    OY 63 ----- ArgSerThrGly 66
    Db 2383 TTCAATGCTATACCTTATATTTTACAAATTTGTTTCTTCTAGTCAAGATCTACTGG 2442
    OY 67 MetValValPheAsnTrpLysAspCysAsnAsnThrLeuGlnLysThrGlu ----- 83
    Db 2443 ATGTAGATTTTCAACTATTAAGATTTGAATATATACATTTACAAAGAACTGAAGGTATGCT 2502
    OY 83 ----- 83
    Db 2503 TTTCGTGTTCTGCACAAATTCGATGTCATGTCATATGTTCTCTAGATGATTTGTTAT 2562
    OY 84 ----- ValArgGluAspCysSerCysProPheCysSe 94

```

Db 2563 TTACTATTTTCTGTATGTCACGAGTTAGGAGGATTTCTGTCATTTTGCTC 2622
 QY 94 rMetLeucylserylserPheLys----- 101
 Db 2623 TATGCTATGTGTACTTCAAGGTGGCACTATTACACTGAGGTTCTCCGGGCTT 2682
 QY 101 ----- 101
 Db 2683 TTCAATCTACACTGTGAATGCTACGCGTTTAATGCTATATCTTCACTGTTCG 2742
 QY 101 ----- 101
 Db 2743 TTACATATTTTGTGTTTGTGTTGCTCTCTTTTAACTGCTAGTGTGTC 2802
 QY 102 -----G 102
 Db 2803 TTATCTGAGAAACACTGTCAGTTCCAGTTCACTTACATCTGCTGTCTATGCAAG 2862
 QY 102 lLeuGlnPheH1sLeuAnsSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
 Db 2863 GCGTCGACATTCATTTGATTCATCTCATGATTTATTGAATTGAGTTCAAGGTATG-T 2921
 QY 121 ----- 121
 Db 2922 GGTTTTANGAATTTCTGTTTGCCATGCGCTTAGTAATGAGGTTATAGTTAAAAAG 2981
 QY 122 -----GluGluTyrGlnThrValAsnValSerValLysLeuA 134
 Db 2982 GGTCTTTCCTATTGTAGCTTTTGGAGAAATACCAAGAGTTAATGTTCTGTAACCTTA 3041
 QY 134 snSerPheL1ePheGlu----- 139
 Db 3042 ATTCCTCATATTGAGGTCACTTAACTTGTTAACTGGAATGGAATCCTATAGCTG 3101
 QY 140 -----GluGluGlySerAspAsp 146
 Db 3102 GTGAAATTTGGTTTATATTCATCTTATTGTACTAGGAAGAAAGATGATGAT 3161
 QY 147 LysPheGluProPheSerLeuCyS----- 154
 Db 3162 AAATTTGAGCCCTTCTCTCTCG-GTAACTCTCAGAACCCCTTGATTAAATACCTTAATA 3220
 QY 154 ----- 154
 Db 3221 GCAGTAACCTCTGCTTTCTTCTGTCAGTACTTCTATAAATCCAAACACATGTTTTCG 3280
 QY 155 ---SerLysProAlaGlyArgArgGlnArgGlyLysArgGlnAsnThrArgArgLeuLys 173
 Db 3281 AGCTCGAAACCTCGTAACCGTAGACAAAGAGGTGGCAAAATTAACACAGGAGACTTAA 3340
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyThr 193
 Db 3341 GTATGCTTTTACCTTGATTCACCCAGTTAGCTAATGCCACGAAATGGAATTTGCC 3400
 QY 194 LeuLeuAsnAsp----- 197
 Db 3401 CTGCGATGATGTGAATAATCACATCTTCTGTGTATGCTGTGGCTTAGAACTTC 3460
 QY 197 ----- 197
 Db 3461 ATTTTACAGAAAGATACAAATGTCCTGATTTGTTTGTACTCTCTCTGCATTC 3520
 QY 198 -----GlyAs 199
 Db 3521 TTCTTGAGGAGTATGTTTACACAGACTGATGTAACAAATTAATGCGATGCTACAGGAA 3580
 QY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetLysSerAs 219
 Db 3581 CCGTGCTTTAGGATATCCGAGGCAACAGACTTCTGAGCAATTTGAGATGACTAGCAA 3640
 QY 219 nIleProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 Db 3641 CATTCACACCACTAGCCCACTCTTCTGTGACGCTGTGCTAAAGTTATATTAACAAC 3700

QY 239 rGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
 Db 3701 CGAAGCTGTGTCTCCCTGCTACTAGACAGAAAGATTATCTGTGAGCATCAGAGGCTAG 3760
 QY 259 9----- 259
 Db 3761 AAGGTTGTTCATCATGACACCCCGTCATCATTAATTACCAATCTGTTTACAAATGTT 3820
 QY 260 -----Se 260
 Db 3821 CTTCCTAATAGATAGTGTGTTTACTGTACTGCCATTAATAACGAGAAATTTCTTCAG 3880
 QY 260 rHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
 Db 3881 CCACCTACTTCTTCGAAACGCCAATTTCTATCATCTCTACAGAGTCCAGTGATCCAAGT 3940
 QY 276 ----- 276
 Db 3941 TCCTTCACCTACTCTCTAGGCATTTTCTTAAATTGCTCANGATGATATCTTATCAAGC 4000
 QY 276 ----- 276
 Db 4001 ATACTGTGTTGTTCTCATCTAAATTTGATTTGATTCTGTATGATCAACGCAAAAA 4060
 QY 276 ----- 276
 Db 4061 ATTATGTCANGTGTCTCCGTTTATTTATGCCACTAACCAAAACCTGATTTCTGTGTA 4120
 QY 277 ---ProMetAlaLeuGlnGluAlaMetSerAspArgAspSerGluAspGluAlaAspAs 295
 Db 4121 CAAGCCAATGCGCGCTTGAGCAAGTAATGTGATGCGATGCGAGATGAGATGAAGTCGATGA 4180
 QY 295 pAspValAlaAspPheGluAspArg----- 303
 Db 4181 CGATGTTGCAGATTTTGAAGATGCGCAGGTATTCATGATTTCTTGTGCTTCATTAAG 4240
 QY 303 ----- 303
 Db 4241 TAGGCAACAGAAATGATATACATGATGTAATGCTTAATGCGCTTTGAAACTTAATAAACG 4300
 QY 304 --GlnMetLeuAspAspPheValAlaValAlaLysAspGluLysGlnPheMetHisLeuT 323
 Db 4301 TGCAGATGCTGATGACTTTGTGGATGGAATTAAGATGAAGAAAGCAATTCATGATCTTT 4360
 QY 323 rPAsnSerPheValArgLysGln----- 330
 Db 4361 GGAAGCTGTTGTAGAAACAAAGTAAGTAATCTTCTGTACACTTGAAACACACAAAA 4420
 QY 330 ----- 330
 Db 4421 AGACCTTATGCTTACATTCATACCTGTCTAAATGATTTGCTTATGGAACCTTGAGCT 4480
 QY 331 -----ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
 Db 4481 CAATTAATGATGATGTTTGCAGGGTTGTAGCAGATGTCATATCTTGTGGCAATGTA 4540
 QY 344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
 Db 4541 GTATTTCAAGATTTTACGAGAAAGATGACGCTGTACTCATCACTCTCTGTAATAT 4600
 QY 361 ----- 361
 Db 4601 AAGTACACCAAAACATATACAGACACATACACTATACATTAATTTTGTTCGTTTTCGA 4660
 QY 362 -----CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisG1 374
 Db 4661 AAGAAATAATAAAAAATCCAGGTGTGGAGATGTTTGTGATTAACATATGGAACCAAG 4720
 QY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
 Db 4721 ACTGTGCACTCAGCCACCATCAACACTGCAATACATCTCTCGAATTTGCGGTAAATAC 4780

[illegible]

Query Match:	8	69.76%	Mismatches:	12	Indels:	577	Gaps:	14
US-09-890-220-2 (1-445) x ATCHRIV45 (1-199548)								
QY	1	MetCysArgGlnAsnCysArgAlaLysSerSerProGluValIleSerThrAspGlu	20					
DB	25399	ATGCTAGGCGAAATTTGTCGGCGGAATCTCACCAGAGAGTGTTCACCTATGAG	25458					
QY	21	AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu	40					
DB	25459	AATCTCTGATATATGTATGTAACCTGTTCGACTATATACATCTTCACCTCGCTCTTA	25518					
QY	41	GlyAsn-----	42					
DB	25519	GCGAA-CGTATGATTTGGCCCTTCCTCATCATCTTTAGCTTAGTAATCTTCATCTCC	25577					
QY	42	-----	42					
DB	25578	TGTGTAGATCACCCACTAATAGTTTGAGTTTGCTAAGCTGATTTATGTGTGACTATGGC	25637					
QY	43	-----Pro	43					
DB	25638	GAGCTGTGCTCTTTTGTCTCCAAATGTTATTTGAACCTGTGTGTTGTTGGACGCA	25697					
QY	44	SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer---	62					
DB	25698	TCGTTTCTGGCAAGATCTTGAACCTACAAAATTGGGGCAAAAGCGCAAAAGAAAGTA-TGC	25756					
QY	62	-----	62					
DB	25757	GTTTCTTCTGAAATGTAGTTGCCACACAGTATATGTTATTTACTTACTTAAATATGAA	25816					
QY	62	-----	62					
DB	25817	GCTATATGACTATTTATCTTTGTTGGAGTAATATGGACAAATAGATGGTTCTCTTTG	25876					
QY	63	-----ArgSerThrGly	66					
DB	25877	TTTCATGCTATACACTTATATTTTACAAAATTTGTTTGTCTAGTGCACAGATCTACTGGG	25936					
QY	67	MetAlaValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu-----	83					
DB	25937	ATGGTAGTTTTCACATATTAAGAGATGTGTAATAATACATTAACAAGAACTAAGGTTAGTCT	25996					
QY	83	-----	83					
DB	25997	TTTTCTGTCTGACAAAATTTGAGATCTCAATGCTATATGTTTCTAGATGATTTGTTAT	26056					
QY	84	-----ValArgLysAspCysSerCysProPheCysSe	94					
DB	26057	TTACTATTTTTTCTGTATGTGTACGACGAGTTAGGGAGATGTTCTTGCCATTTTGGCTC	26116					
QY	94	rMetLeuCysGlySerPheLys-----	101					
DB	26117	TATGCTATGTGTAGTCTCAAGGTGGCAACTATTAACAAGTGAAGTTCTTCGCGGGGCT	26176					
QY	101	-----	101					
DB	26177	TTTCATATCTAACACTGTGAATGTCTACTGCCGTTTAATGCTATATACTTTCACCTGTTGG	26236					
QY	101	-----	101					
DB	26237	TTTACATATTTTTTGTGTTTGTGTTTGTCTTGTCTTTTAACTGCGTAGTGTGTC	26296					
QY	102	-----G	102					
DB	26297	TTATCTAGAAACATGTCTCCAGTCCAGCTTACAAATCCATGTGCTTGTGCTATACAGG	26356					
QY	102	LyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGlnPheLysLeuPhe-	121					
DB	26357	GGCGCAATTTTCATTTGAATTCATCTCATATTTATTTTGAATTTGAGTTCAAGGTATG-T	26415					

```

OY 121 ----- 121
Db 26416 GGTATTATGAATTTCTGTTTTCCTATGCCGTTAGTATAGAGTTATAGTAAAG 26475
OY 122 ----- GluGluTyrGluThrValAsnValSerValLysLeu 134
Db 26476 GGTCTTCCATGTAGCTTTGGAAAGATACAGACAGTTATGTCTGTAAACTTA 26535
OY 134 snSerPheIlePheGlu----- 139
Db 26536 ATTCTTCATATTGAGGTCAGTACTTAAACTTGTTAAATGGGAAATCCTATAGCTG 26595
OY 140 ----- GluGluGlySerAspAsp 146
Db 26596 GTGAAATTTGGTTTATATTCATCCTTATTTGTACTAGAGAGAAAGATGATGAT 26655
OY 147 LysPheGluProPheSerLeuCys----- 154
Db 26656 AAATTGAGCCCTTCTCTCTG-GTAACTCTAGAACCCCTGTAAATACCTTAATA 26714
OY 154 ----- 154
Db 26715 GCAGTAACCTCTGCTTTCTGTCAGTACTTCTCATTAATCCAAACCAATGTTTGC 26774
OY 155 ---SerLysProArgLysArgArgLysArgLysArgLysArgLysArgLys 173
Db 26775 AGCTCGAAACCTCTGAAGCGTAGACAAAGAGTGACAAATTAACACAGAGACTTAA 26834
OY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyThr 193
Db 26835 GATACGCTTTTACCGTTGGATTCACCCATTTAGCTAATGACAGAAATAGATGCC 26894
OY 194 LeuLeuAspAsp----- 197
Db 26895 CTGCTGAATATGTGTAATACACATCTTCTGTGGTATTCGTTGGCTTACAACTTC 26954
OY 197 ----- 197
Db 26955 ATTTTACAGAAAGATACAAATGCTCATGTTGTTTGTACTTCTCTGCATTC 27014
OY 198 ----- GlyAs 199
Db 27015 TTCTTGAGGGAATATTACCAAGACTGATGTACAAATTAATGCTGTACAGGAAA 27074
OY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetTrpSerAs 219
Db 27075 CCGTGGTTTGGATATCCGAGGCAACAGAGCTGTGACAAATTTGAGATGACTACAA 27134
OY 219 nileProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuTrpSe 239
Db 27135 CATTCACAGCCATAGCCCATCTCTCTGAGCGCTGTGCTAAAGTTATATTAAACAC 27194
OY 239 rgLualAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluAlaTrpSerGluAlaTr 259
Db 27195 CGAAGCTGTGGTCCCGTACTAAGACAAAGATATCTGCTGAGCATCAGAGGCTAG 27254
OY 259 g----- 259
Db 27255 AAGGTTTGTTCATCAGACACCCCGCATCATTAATACCATACCTGTTGTACAAATGTT 27314
OY 260 -----Se 260
Db 27315 CTTCCTATTATGATAGTGTACTGTACTGCATTAATTAACCGAGAAATTTCTTCAG 27374
OY 260 rHisLeuLeuLeuGlnLysArgLysPheTyrHisSerHisArgValGln----- 276
Db 27375 CCACCTACTCTTCAGAAAGCCCAATTCTATCATCTCACAGAGCCAGGTGATCAAGT 27434
OY 276 ----- 276
Db 27435 TCCTGACCTACTTCTTAGCATTTTCTTAATTTGTCATGATGATGATCTTATCAAGC 27494

```

```

OY 276 ----- 276
Db 27495 ATACTGTGGTTTCTCATCTAAATTTGTATTTGATCTGATGATTCACGCAAAAA 27554
OY 276 ----- 276
Db 27555 ATTAGTCCAGTCTCTCCCTTTTATTCGCACTAACCAAAAACGCAATGTTCTTGGA 27614
OY 277 ---PrometAlaLeuGluGlnValMetSerAspArgAspSerGluAspValAspAs 295
Db 27615 CAAGCCAAATGGCGCTTGAGCAAGTAATGTCTGATCGGATGAGGAGATGAAGTCATGA 27674
OY 295 pasPValAlaaspPheGluaspArg----- 303
Db 27675 CGATGTGCAGATTTTGAAGATCCCGAGATTCATGATTTCTTCTCGTTCAATTAAG 27734
OY 303 ----- 303
Db 27735 TAGCCACAGAAAATGATAGATGATGATGATGATGATGATGATGATGATGATGAT 27794
OY 304 ---GluMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeu 323
Db 27795 TCACATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 27854
OY 323 rPasSerPheValArgLysGln----- 330
Db 27855 GGAACGCTTTGTAGAAACAAAGGTAATGATGATGATGATGATGATGATGATGATGAT 27914
OY 330 ----- 330
Db 27915 AGACCTTATGCTTACATTCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 27974
OY 331 ---ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
Db 27975 CAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28034
OY 344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
Db 28035 GATTTTCAAGATTTTACAGAGAAAGAGTCTGATGATGATGATGATGATGATGATGAT 28094
OY 361 ----- 361
Db 28095 AAGTACACCAAAATATACAGACATTAATGATGATGATGATGATGATGATGATGATGAT 28154
OY 362 ---CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlu 374
Db 28155 AAGAAAATTAATAAATTCACAGGTGTGATGATGATGATGATGATGATGATGATGAT 28214
OY 374 yLeuValAspSerAlaThrIleAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 394
Db 28215 ACTGTGACACACGACACACACACACACACACACACACACACACACACACACACACAC 28274
OY 394 rSerAspThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 414
Db 28275 CTCA-----GTCACTAACACACACACACACACACACACACACACACACACAC 28338
OY 414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnValAspAs 434
Db 28329 CAACACACATTAACATTTGGATCTCCGAAATGACATTAATAAACAAAGACAAATGTTGACAA 28388
OY 434 nLysAspAsnAsnSerArgAspLys 442
Db 28389 CAAGCACAAATTAACAGACAGACAG 28413
OY 28389 CAAGCACAAATTAACAGACAGACAG 28413

```

RESULT 15
 AFPCAT 201471 bp DNA linear PLN 29-JUN-1999
 LOCUS Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment
 DEFINITION No. 7.
 ACCESSION 297342
 VERSION 297342.2 GI:5302800
 KEYWORDS SOURCE Arabidopsis thaliana.

ORGANISM	Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.	
REFERENCE	1 (bases 1 to 201471)
AUTHORS	Bevan, M., Stiekema, W., Murphy, G., Wambutt, R., Pohl, T., Terry, N., Kreis, M., Kavanagh, T., Entlan, K.D., Rieger, M., James, R., Puidomenech, P., Hatzopoulos, P., Obermayer, B., Duesterhoff, A., Jones, J., Palme, K., Ansgore, W., Delseny, M., Bancroft, I., Mewes, H.W., Scheller, C. and Chalhatzis, N.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 201471)
AUTHORS	EU Arabidopsis sequencing project.
TITLE	Direct Submission
JOURNAL	Submitted (25-JUN-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopfersplitz 18a, D-82152 Martinsried, FRG, E-mail: schnelleimips.biochem.mpg.de,mayeremips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT	On Jun 30, 1999 this sequence version replaced gi:2245031. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATFCa6 at the 5' end and an overlap with ATFCa8 at the 3' end.
FEATURES	Location/Qualifiers
source	1..201471
	/organism="Arabidopsis thaliana"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
source	1..8956
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid EM16B10.IB5K23"
source	21108..37844
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid EM16B10.IB6C5"
source	37120..39085
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid EM16B10.IB4A23"
source	37904..128469
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="BAC IGF5D3"
source	38368..63818
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid CC3M3"
source	106243..128469
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid CC16N19"
source	113141..132730
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid CC20I2"
source	129847..157017
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid CC1B24"
source	152677..190349
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid G16599"
source	185303..201471
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="cosmid G11319"
gene	12153..13466
	/gene="d14415w"
	/join(12153..13312,12416..12714,12983..13062,13151..13466)
	/gene="d14415w"
	/note="contains EST gb: 1216777,935272;906153;947895;315626"
	/codon_start=1
	/product="DNA-binding homeotic protein Atb1-2"
	/protein_id="CAB10452.1"
	/db_xref="GI:2245033"
	/db_xref="SWISS-PROT:005466"
	/translation="MWFEDDVLGLSLGLNFPKKQINLKNPSVSYTPSSSGSGLEFRRS
	SWNESTSSVPNSDSQKERTFIRGIDVNRPSSTAEYGEDADGAVSSPSTYSSSTGK
	RSEREDTDPQSGRIGIDDEDGNSRKLRLSDSAILEETFKDSTLNPQOKOLA
	KQLGLARQVEVWFQNRARATKIKQTEVCEFLRRCCELTLEENRLQREVELRALK
	LSQRYMHSPPPTITLMPCSCHEVSPPPQQAATSAHRSILPVANMAPATISGLT
	FPALPRRS"
exon	12153..12312
	/gene="d14415w"
	/number=1
Intron	12313..12415
	/gene="d14415w"
	/number=1
exon	12416..12714
	/gene="d14415w"
	/number=2
Intron	12715..12982
	/gene="d14415w"
	/number=2
exon	12983..13062
	/gene="d14415w"
	/number=3
Intron	13063..13150
	/gene="d14415w"
	/number=3
exon	13151..13466
	/gene="d14415w"
	/number=4
gene	14609..16030
	/gene="d14420c"
	/complement(14609..16030)
gene	/gene="d14420c"
	/complement(14609..16030)
CDS	/note="sequence differences between cDNA and gDNA similarity to hydroxyproline-rich glycoprotein (clone HYP3.6) kidney bean (fragment)"
	/codon_start=1
	/product="glycoprotein homolog"
	/protein_id="CAB46042.1"
	/db_xref="GI:5302801"
	/translation="WEARSIKKPTIOLGKEDONPKKFSRFTFKALITLVICAVPV
	FISQPELANQTRLELHLVYGVTAIVSGLSRRRYDGGGGGTSNSHNRKADSHN
	NSHSTVPKILVSVFNVGHESESEPSDUSGQQRFFQTKKNTKHKTIPEVTRFVDR
	VSEENNEKPLLPVRLNLSRVSDSGNSGNGWKKVRSRELKTLGDGNSDVLPSPI
	PWRSRSSSSSSSSKVESLPSVKNTLVESDPLKNTLPPSPSSFSRPSNIPNLAS
	EPHPSPPPPPPPPLPPLAVNGSSRKRDGIVRVERESSVHKTRKAGSGPFPPPP
	PPPPVYVYVSPPTKPRLSNERKSSSQKKRRAPKVKVMSDPLVESKEQDTKQORS
	NIGSKAVESNGEQRGNEIHDEVEKTIVEEGVSEINNSDDVKKADEFLAKFRE
	QIRLRIEISIKRSTNKISANSSR"
exon	/complement(14609..16030)
	/gene="d14420c"
	/number=1
repeat_region	16148..16170
	/note="23bp T tandem repeat"
misc_feature	17691..19115
	/note="CDNA1KMY"
gene	17793..19171
	/gene="d14425c"
	/complement(join(17793..18046,18147..18250,18348..18399,
	18481..18571,18886..18960,19058..19171))
	/gene="d14425c"
	/complement(join(17793..18046,18147..18250,18348..18399,
	18481..18571,18886..18960,19058..19171))
CDS	

QY 154 ----- 154
 Db 40886 GCAGTAACCTCGCTTTCTTGTGACACTCTCTATTAATCCACACACATGTTTGC 40945
 QY 155 ---SerLysProAlaGlySerArgGlnArgGlyGlyArgGlnAsnThrArgArgLeuLys 173
 Db 40946 AGCTCGAAACCTCGTAACGCTAGACAAAAGAGTGGCAGAAATTAACCCAGCGAGACTTAAA 41005
 QY 174 ValLysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyLethr 193
 Db 41006 GTATGCTTTTAACTGCTGGATTCACCCAGTTTGTAGCTAATGGCAGAAAATGGAATTC 41065
 QY 194 LeuLeuAsnAsp----- 197
 Db 41066 CTGGCAATGATGTAATAATACACTCTCTCTGCTGATTCGTTGCGCTTAGAACCTTC 41125
 QY 197 ----- 197
 Db 41126 ATTTTACAGAAAGATACAAATGCTCGATGTTTGTAGTTTGTACTCTCTCGCATTC 41185
 QY 198 -----GlyAs 199
 Db 41186 TTCTTTCGAGGCTAATGTTTACCAGACTGATGTACAAAATTAATGGCATGCTACAGAAA 41245
 QY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetLthrSerAs 219
 Db 41246 CGTGTGTTTAGGATATCCCGAGCGACAGAGCTTCTGACACTTTGAGATGACTAGCAA 41305
 QY 219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 Db 41306 CATTCACACGACCAATAGCCCACTCTCTGACGCTGGTGTAAAGTTATATTAACAC 41365
 QY 239 rGluAlaValAlaProAlaThrIleThrArgLysLeuSerIleGluLysSerGluAlaAr 259
 Db 41366 CGAACTGTGGTCCCTGCTACTAGACAGAAAGTTATCTGTGCGCATGAGAGCTTAG 41425
 QY 259 g----- 259
 Db 41426 AAGGTTGTTCATCATGACACCCCGCTCATCATTAATTAACATACCTGTTTACAAATGT 41485
 QY 260 -----Se 260
 Db 41486 CTTCCTATTTATGATAGTATGTTTACTGTACTGCCATATTAACCGAGAAAATTTCTTCAG 41545
 QY 260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
 Db 41546 CCACCTACTTCTTGAGAAACGCCAATTTATCATCTTCACAGAGTCCAGGATCCAAAGT 41605
 QY 276 ----- 276
 Db 41606 TCCTTCACCTACTCTTAGGCATTTTCTTAATTGTCTCATGATGATATCTTATCAAGC 41665
 QY 276 ATACTGTGTTGTCTCATCTAAATTTGTATTTGATTCTGTATGATCAACGCAAAAAA 41725
 Db 41726 ATATATGTCATGTTGTCTCCGTTTATTTATGTCACCTAACCAAAAACGTCATGTTTCTTGCA 41785
 QY 277 ---PrometaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
 Db 41786 CAAGCATATGGCGCTTAGAGCAAGTAATGTGATCGGATAGCGAGATGAAGTGCATGA 41845
 QY 295 PAspValAlaAspPheGlnAspArg----- 303
 Db 41846 CGATGTTCAGATTTTGAAGTCGCCAGATATTCATGATTTCTTTCTGCGTTCAATTAAG 41905
 QY 303 ----- 303
 Db 41906 TAGGCACAGAAATGATATACGATGTACTTGTCTTAATGGCTTTTGAACACTTAAAAAAAGC 41965
 QY 304 --GlnMetLeuAspAspPheValAspValAsnLysAspGlnPheMetHisLeuT 323

Db 41966 TGCAGATGCTGATGACTTTGGGATGGAAATGAAGAAAGCAATTCATGACTTTT 42025
 QY 323 rPAsnSerPheValArgLysGln----- 330
 Db 42026 GGAACCTGTTTGTAAAGAAACAAAGTAACACTTCTCTTACACTTGAACACACAAAA 42085
 QY 330 ----- 330
 Db 42086 AGACCTTATGCTTACATTCATTCATACCTGCTTAATGATTCGTTATGAAACTTGGACT 42145
 QY 331 -----ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
 Db 42146 CAATTAATGATGATGTTTGCAGGGTTATAGAGATGGTCATATCTTGGGCATGTGA 42205
 QY 344 AlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
 Db 42206 GTATTTCAAGATTTTACAGAAAGAGTTGCACTGTTACTCATCACTCTCTGGTAATAT 42265
 QY 361 ----- 361
 Db 42266 AAGTACACCAACATATACAGACATTAACACTATCAATTTTGTTCGTTTCTGA 42325
 QY 362 -----CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGly 374
 Db 42326 AAGAAATTAATAAATTCAGAGTGTGGAGATGTTTGTATTAACATATGAACCATGG 42385
 QY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
 Db 42386 ACTTGTGACATCAGCCACCATCAACACTGCATACCTCCGAGAAATTTGCCGTAAATAC 42445
 QY 394 rSerAspThrTrpThrTrpAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
 Db 42446 CTCA-----GTCATTAACACACACACACAGCTGTGGATCATCCAGTACTCAAAAC 42499
 QY 414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
 Db 42500 CAACAAATTAACATTTGGATGATCCGATGACATTAATAAACCAAGCAATGTTGACAA 42559
 QY 434 nLysAspAsnAsnSerArgAspLys 442
 Db 42560 CAAGGACAAATAACAGCAGAGACAAAG 42584

Search completed: June 20, 2003, 00:55:20
 Job time : 3851 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 22:08:48 ; Search time 330 Seconds

(without alignments)
3036.786 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378
Sequence: 1 MCRONCRAKSSPEEYISDNE.....INKNKNVNDKNNSDXYIK 445

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q/cgu2.1/USPTO_SPOOL/US09890220/runat_13062003_144850_16438/app_query.fasta.1.583
-DB=N.Geneseq_101002 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220.ecgn1.1.263.ernat_13062003_144850_16438 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2378	100.0	1722	21	AAA63670	Nucleotide sequence
2	2378	100.0	1722	21	AAA47751	VNR2 nucleic acid.
3	2285	96.1	1715	21	AAA63671	Nucleotide sequence
4	2285	95.0	1715	21	AAA47753	VNR2 nucleic acid.
5	2258.5	95.0	1737	21	AAA47755	Aberrantly spliced
6	1923	80.9	1497	21	AAAC51581	Arabidopsis thaliana
7	1813.5	76.3	6338	21	AAA47752	VNR2 nucleic acid.
8	1659	69.8	5895	21	AAA47754	VNR2 nucleic acid.
9	923	38.8	2280	21	AAZ36947	CDNA encoding a MP
10	906.5	38.1	2248	21	AAZ36953	CDNA encoding an O
11	451.5	19.0	558	21	AAA47758	At Hyp 2245035 (AT
12	407.5	17.1	5580	21	AAZ36948	DNA encoding a MPC
13	368.5	15.5	325	21	AAA47757	Arabidopsis thaliana
14	340.5	14.3	640	21	AAAC44147	Arabidopsis thaliana
15	327	13.8	319	21	AAAC51582	Arabidopsis thaliana
16	324	13.6	525	21	AAAC54919	Arabidopsis thaliana
17	311	13.1	463	21	AAA47756	Rice expressed seq
18	225	9.5	3545	22	AAZ52965	Human CDNA encodin
19	221	9.3	4441	21	AAA47759	KIRA0160 CDNA. Ho
20	221	9.3	4441	21	AAZ52965	Pancreas cancer re
21	219	9.2	3572	23	ABL29637	Drosophila melanog
22	208	8.7	7021	23	ABL29636	Drosophila melanog
23	196.5	8.3	1216	21	AAZ36962	DNA fragment encod
24	157.5	6.6	1257	24	AAZ52945	CDNA sequence #428
25	153	6.4	2481	22	AAZ52945	Human CDNA encodin
26	131.5	5.5	3738	21	AAZ70178	Plasmodium falcipa
27	128.5	5.4	5454	21	AAZ70236	Plasmodium falcipa
28	127.5	5.4	3300	21	AAZ52945	DNA encoding yeast
29	127	5.3	2466	23	ABL27385	Drosophila melanog
30	127	5.3	8045	21	AAA65171	Neurospora crassa
31	126	5.3	646	24	AAZ46952	Oligonucleotide fo
32	126	5.3	646	24	ABQ46953	Oligonucleotide fo
33	124.5	5.2	2061	21	AAZ70228	Plasmodium falcipa
34	124	5.2	557	24	ABQ43458	Oligonucleotide fo
35	124	5.2	557	24	ABQ43459	Oligonucleotide fo
36	124	5.2	5340	22	AAZ6289	P. falciparum egfs
37	123.5	5.2	6033	21	AAZ70152	Plasmodium falcipa
38	122.5	5.2	4311	21	AAZ70133	Plasmodium falcipa
39	122	5.1	5145	21	AAZ70209	Plasmodium falcipa
40	121.5	5.1	412	21	ABN81085	Shrimp polynucleot
41	120.5	5.1	198	17	AAZ17204	DNA-spanner oligon
42	120.5	5.1	198	17	AAZ17205	DNA-spanner oligon
43	120	5.0	7143	21	AAZ70250	Plasmodium falcipa
44	119.5	5.0	567	21	AAZ29550	HIV codon altered
45	119	5.0	3282	24	ABN79865	Fungal ZRC gene se

ALIGNMENTS

RESULT 1
ID AAA63670 standard; CDNA; 1722 BP.

AC AAA63670;
XX
DT 04-DEC-2000 (first entry)

DE Nucleotide sequence of VNR2 sequence of Arabidopsis sp.

XX H51: one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; VNR2; ss.

OS Arabidopsis sp.

XX WO200046358-A2.

XX 10-AUG-2000.

XX 25-JAN-2000; 2000MO-GB00197.
 PF 05-FEB-1999; 99GB-0002660.
 XX (PLAN-) PLANT BIOSCIENCE LTD.
 PA Johnson U, West J, Dean C;
 FI WPI; 2000-532899/48.
 DR
 XX New nucleic acid derived from the FRI locus of a plant, e.g.
 PT Arabidopsis, encoding a polypeptide capable of specifically altering
 PT the flowering time of a plant -
 XX
 PS Example 5; Page 53-54; 73pp; English.
 CC The present sequence represents a VRN2 sequence from variety Landsberg
 CC erecta. The specification describes a sequence which encodes a
 CC polypeptide capable of specifically altering the flowering time of a
 CC plant. The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
 CC of Arabidopsis. The FRI polynucleotide is used to transform plants, so
 CC that the flowering time of a plant is altered. This is used, for example,
 CC for plants in which the leaves or tubers are a commercial product, where
 CC it is desirable to avoid 'blotting' (initiation of flowers and stem
 CC elongation) at too early a stage. Conversely, it may be desirable to
 CC alter flowering under certain circumstances e.g. to vary flower
 CC production across the seasons.
 XX
 SQ Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,53e-243 Length: 1722
 Score: 2378.00 Matches: 445
 Percent Similarity: 100.008 Conservative: 0
 Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 21 Indels: 0
 DB: Gaps: 0
 US-09-890-220-2 (1-445) x AAA63670 (1-1722)
 QY 1 MetCysArgGlnAsnCysArgAlaValSerSerProGluGluValIleSerThrAspGlu 20
 DB 231 ATGTGTAGCGAGATGTCGCCGAATCTCACCGAGAGATGATTCAACGATGAG 290
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 291 AATCTCTTGATATATGTAAACCTGTCACATAATACATCTTCACCTTCGCTCTTA 350
 QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 351 GGCACACCATCGTTCTCCCAAGATGCTTGAACTGCAAAATTTGGAGCAACGCCAAGA 410
 QY 61 LysSerArgSerThrGlyMetValIlePheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 DB 411 AAGTCAGAGATCTAGGATGATGATGTTTCACACTAATAGATTTGAATTAACACTTCAG 470
 QY 81 LysThrGluValArgGluAspCysSerPhePheCysSerMetLeuCysGlySerPhe 100
 DB 471 AAAACTGAAGTGGAGAGATGTTCTTCTTCATTTTGGCTCATCTATGCTGAGCTTC 530
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
 DB 531 AAGGGCTGCAATTTCAATTCATTCATCATGATTAATTAATTCCTCATATTTGAGAGAA 590
 QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 591 TTGCAAGAAATACAGACAGATTAATGTTCTGTAATAACTTAATTCCTCATATTTGAGAGAA 650
 QY 141 GluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 DB 651 GAAAGAGATGATGACATTAATTTGAGCCCTCTCTCTGCTGCAAACTCGTAAGCGG 710

QY 161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValLysPheLeuProLeuAsp 180
 DB 711 AGACAAAGAGGTGGACAAATTAACACCGAGAGACTTAATGATGCTTTTACCGTGGAT 770
 QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
 DB 771 TCACCCAGTTTAACATAAGGCAGAAAATGAATCCCTACTTATATATGAATGAACCTG 830
 QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyIlePheGluMetThrSerAsnIle 220
 DB 831 GGTTAGGATATCCCGAGCAACAGAGCTTCCTGGCAAAATTTAGATGACAGCAACAT 890
 QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 DB 891 CCACAGCCATAGCCCACTCTCTCTGACGCTGCTTAAGTTATTTGACAAAGCGAA 950
 QY 241 AlaValAlaProAlaThrIleThrArgLysLeuSerIleGluArgSerGluAlaArgSer 260
 DB 951 GCTGTGTCCTCGTACTTAAGACAAAGATTAATCTGAGCGATCAGAGGCTAGAACG 1010
 QY 261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 DB 1011 CACCTACTCTTCAGAAAAGCCCAATGCTATCATCTTCACAGAGTCCAGGCATAGGCGCTT 1070
 QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
 DB 1071 GAGCAAGTAAATCTCGACCGGATAGCGAGATGACATGCAATGTCATGTCACATTTT 1130
 QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 DB 1131 GAGAGATCCGACATGCTTGATGACTTGTGTGATGTGATTAAGTAAAGCAATTCATG 1190
 QY 321 HisLeuTyrPasnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
 DB 1191 CATCTTGGAACTCGTTTGTAAGAAACAAAGGTTTATGACATGTCATATCTCTTGG 1250
 QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
 DB 1251 GCATGTGAAGCATTTTCAAGATTTTACAGAAAGAGTTGCCCGTTACCATCATCTCTTC 1310
 QY 361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
 DB 1311 TGGTGTGGAAGTGTGTTTGTGTAATTAACATGAGACCATGCACTTGCACACGACAC 1370
 QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgGlnSerSerAspThrThrThr 400
 DB 1371 ATCAACAACCTGCAATACCATCTCGAGAAATTTGCCGTAAATGCTCAGACACACAC 1430
 QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 DB 1431 AACCAACCAACAGTGTGATGATCGTCCAGAGATCAACCAACCAACATTAACATGTC 1490
 QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
 DB 1491 GATCATCCCATGACATTAACAAACAAATACATTTTGACAAAGACATTAACAGCAGA 1550
 QY 441 AspLysValIleLys 445
 DB 1551 GACCAAGTAAATTTAAA 1565
 RESULT 2
 AAA47751
 ID AAA47751 standard; cDNA; 1722 BP.
 XX AAA47751;
 AC
 XX 16-NOV-2000 (first entry)
 DT
 XX
 XX VRN2 nucleic acid.
 DE
 XX
 XX VRN2 gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ss.

```

XX Arabidopsis thaliana var 'Landsberg erecta'
OS
XX Key Location/Qualifiers
FH 231..1568
FT CDS
FT
FT mutation
FT /product= "VRN2 polypeptide"
FT replace(1199,A)
FT /tag= b
FT /note= "found in mutant vrn2-1; alters a Trp codon
FT to a stop codon leading to a truncation of
FT the protein"
XX
XX MO200044918-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000MO-GB00248.
XX
XX 28-JAN-1999; 99GB-0001927.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Dean C, Gendall A;
XX
XX WPI: 2000-499333/44.
XX
XX P-PSDB: AAB00060.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
XX PT with altered vernalization response, flowering time, leaf size and/or
XX PR shape or shade avoidance response for maximized reproductive success
XX
XX Claim 4; Fig 6; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX CC plant encode polypeptides which are capable of affecting one or more
XX CC vernalization responses such as, flowering time, leaf size and/or
XX CC shape or the shade avoidance response of a plant into which the
XX CC nucleic acid is introduced. Introducing such sequences into plants
XX CC to alter these characteristics maximises the reproductive success of
XX CC the plant.
XX
XX SQ Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;
XX
Alignment Scores:
Pred. No.: 3 53e-243 Length: 1722
Score: 2378.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-890-220-2 (1-445) x AAA47751 (1-1722)
OY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 231 AGGTGTAGGCAAAATGTCGCCGAATCTCACGAGAGATGATTCAACTGATGAG 290
OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 291 AATCTCTGTATATATGTTAAACCTGTCACATATATAAACATCTTCACCTTGCTCTCTA 350
OY 41 GluAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 351 GGCACACCATCTTTCTTCCAGATGCTTGAACCTCAAAATGAGCAAGGCAAAAGAA 410
OY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 411 AAGTCAAGATCTACTGGGAGTGTAGTTCACACTTAAGAGATTGTATACACATTACAG 470
OY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 471 AAAACTGAAGTTAGGAGGATGTTCTTGTCCATTTTGTCTATGCTATGTGTAGCTTC 530

```

```

OY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 531 AAGGGCTGCAATTCATTTGATTAATTCATCTCATGATTATTTGAATTTAGTCAAGCT 590
OY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 591 TTCGAAGAAATACACAGACTTAATGTTTCTGTAATAAAGCTTAATCCCTCATATTGAGAA 650
OY 141 GluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 651 GAAGAAAGTGAAGAGATTAATTTGAGCCCTCTCTCTCTGCTCCAAACCTCGTAAGCG 710
OY 161 ArgGlnArgGlyLysValArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 711 AGACAAAGAGGTGGCAAAATACACACAGAGACTTAAGATGCTTTTACCGTTGGAT 770
OY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
DB 771 TCACCCAGTTTAACATAATGACACAGAAATGGAATCACCTTAATGATGAAACCGT 830
OY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 831 GGTTAGGATATCCCGAGGCAACAGAGCTTGTGACAAATTTGAGATGCCAGCAACATT 890
OY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 891 CCACCAAGCATAGCCCACTTCTCTGAGCGCTGGGTAAAGTTATATTGACAAACGGA 950
OY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 951 GCTGTGGTCCCTGCTACTAGAACAAAGAAATGTTATGCTGAGAGATCAAGGCTAGAAC 1010
OY 261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1011 CACCTTAATCTTCAGAAAGCCCAATTCATCTTCACAGAGTCCAGCAATGGCGCT 1070
OY 281 GlnGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
DB 1071 GAGCAAGTAAATGTCTGACGGGATGACGAGATGAAGTGAATGATGACATGTTCCAGATT 1130
OY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspLysGlnPheMet 320
DB 1131 GAAGATCGCCAGATGCTTGATGATCTTGTGATGTGAATGAACCATGAAAGCAATTCATG 1190
OY 321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB 1191 CATCTTTGAACTCGTTGTGAAGAAACAAAGGTTATGACAGATGCTATATCTTGG 1250
OY 341 AlaCysGluAlaPheSerArgPheTyrGlyLysGlyLeuHisArgTyrSerSerLeuPhe 360
DB 1251 GCATGTGAAGCATTTTCAAGATTTTACGAAAGAGCTTCCACCGTTACTCATCTCTTC 1310
OY 361 TrpCysTrpArgLeuPheLeuIleLysLeuThrPasnHisGlyLeuValAspSerAlaThr 380
DB 1311 TGGTGTGAGATGTTGTTTGAATTAACATATGAACCATGAGACTTGTACACAGCACCC 1370
OY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
DB 1371 ATCAACAACTGCAAATACATCTCTCGAATTTCCGTAATTAAGCTCAACACACACACAC 1430
OY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAspThrAsnAsnAsnIleVal 420
DB 1431 AACCAACACACACAGTGTGATGCTCCAGTCACTCAACACACACACATTAATATTGTC 1490
OY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
DB 1491 GATCATCCCAATAGACATAAACAAGAAATGTTGACAAACAGACAAATTAACAGCAGA 1550
OY 441 AspLysValIleLys 445
DB 1551 GACAAAGTAATTTAA 1565

```

```

RESULT 3
AAA63671
ID AAA63671 standard; cDNA; 1715 BP.
XX
XX AAA63671:
AC
XX
XX 04-DEC-2000 (first entry)
XX
XX Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
DE
XX
XX H51: one locus-FRIGIDA; FRI gene; flowering time; blotting;
XX flower initiation; stem elongation; flower production; VRN2; ss.
XX
XX Arabidopsis sp.
OS
XX WO200046358-A2.
XX
XX 10-AUG-2000.
XX
XX 25-JAN-2000; 2000MO-GH00197.
XX
XX 05-FEB-1999; 99GB-0002660.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Johanson U, West J, Dean C;
XX
XX WPI: 2000-532899/48.
XX
XX New nucleic acid derived from the FRI locus of a plant, e.g.
XX Arabidopsis, encoding a polypeptide capable of specifically altering
XX the flowering time of a plant.
XX
XX Example 5: Page 54-55; 73pp: English.
XX
XX The present sequence represents a VRN2 sequence from variety Columbia.
XX The specification describes a sequence which encodes a polypeptide
XX capable of specifically altering the flowering time of a plant.
XX The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
XX of Arabidopsis. The FRI polynucleotide is used to transform plants, so
XX that the flowering time of a plant is altered. This is used, for example,
XX for plants in which the leaves or tubers are a commercial product, where
XX it is desirable to avoid 'blotting' (initiation of flowers and stem
XX elongation) at too early a stage. Conversely, it may be desirable to
XX alter flowering under certain circumstances e.g. to vary flower
XX production across the seasons.
XX
XX Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 2,88e-233 Length: 1715
Score: 2285.00 Matches: 428
Percent Similarity: 97.51% Conservative: 3
Best Local Similarity: 96.83% Mismatches: 9
Query Match: 21 Indels: 2
DB: 1 Gaps: 1
US-09-890-220-2 (1-445) x AAA63671 (1-1715)
OY 1 MetCysArgGlnAsnGlySerProGluGluValIleSerThrAspGlu 20
DB 233 ATGCTAGGAGATGTCGCGAAATCTCAGCGAGAGATGATTTCACATGATGAG 292
OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 293 AATCTCTTGATATATGTAACCTGTCACATATATACATCTTTCACCTCGCTCTCTA 352
OY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrIleGlyAlaIleArgLysArg 60
DB 353 GGCACACCATCGCTTTCGCAAGATGCTTGAACATCAAAATTTGGGCAAAACGCAAAAGA 412
OY 61 LysSerArgSerThrGlyMetValIlePheAsnTyrLysAspCysAsnAspThrLeuGln 80

```

```

DB 413 AAGTCACAGATCTACTGGGATGGTATGTTTCAACTATATAGAGATTGTAATATACATTACAA 472
OY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 473 AGAAGTGAAGTGAAGGAGATGTTGTTGTCATATTTGCTATGCTATGCTGATGCTTC 532
OY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGlyLeu 120
DB 533 AAGGGCGTCGACATTTTCATTGATATCATCTCATGATTTATTAAATTTAGTTCAAGCTT 592
OY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGln 140
DB 593 TTGGAAGAAATACACAGACGTTAAAGTTTGTAAACTTAAATCCCTTCATATTTGAGGAA 652
OY 141 GluGlySerAspAspAspPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 653 GAAGGAGAGATGATGATTAATTTGAGCCCTCTCTCTGCTGCAAACTCGTAAGCGT 712
OY 161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValLysPheLeuProLeuAsp 180
DB 713 AGACAAAGAGGTGGGAGAAATATACACAGAGACTTAAGTATGCTTTTACGTTGAT 772
OY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIlePheLeuAsnAspGlyAsnArg 200
DB 773 TCACCCAGTTTATGCTAATGCGACAGAAATGGAATTCCTGCTGAATATGTAACCT 832
OY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 833 GGTTAGGATATCCCGAGGACAGAGCTTGTGACAAATTTGAGATGACTGACAACTT 892
OY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 893 CCACCAAGCATAGCCCACTCTTCTGAGAGCTGGGCTGAATATTAACAACCGAA 952
OY 241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 953 GCTGTGGTCCCTGTACTACAGCAAGAACTTATCTGTCAGGACAGCAAGCTGTGAAGC 1012
OY 261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1013 CACCTACTCTTTCAGAAAGCCCAATCTATCATCTTCACAGATCCAGCAATGCGCTT 1072
OY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
DB 1073 GAGCAAGTAAATGCTGATCGGAGATGCGAGATGAATGATGATGATGATGATGATTTT 1132
OY 301 GluAspArgGlnMetLeuAspAspPheValAspValAspLysAspGluLysGlnPheMet 320
DB 1133 GAAAGATGCCAGATGCTTGTGATGACTTGTGATGATGATGATGATGATGATGATGATG 1192
OY 321 HisLeuThrAsnSerPheValArgGlyGlnArgValIleAlaAspGlyHisIleSerThr 340
DB 1193 CATCTTTGGAAGCTGTTGTGAGAAACAAAGGTTTATGACAGATGCTATATCTCTTGG 1252
OY 341 AlaCysGluAlaPheSerArgPheTyrGlyGluGlyLeuHisArgTyrSerSerLeuPhe 360
DB 1253 GCATGTGAAGTATTTTCAAGATTTTACGAGAAAGATGACGCTTATCATCATCTTTC 1312
OY 361 TrpCysThrArgLeuPheLeuIleLysLeuThrPheHisGlyLeuValAspSerAlaThr 380
DB 1313 TGGTGTGAGAAATGTTTGTGATTAATCAATGAGAACCATGATGATGATGATGATGATGAT 1372
OY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
DB 1373 ATCAACAACATCAATACCATCTTCGAGAAATTTGCCGTAATTAACCTCA-----GTCACTAAC 1426
OY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAspThrAspAsnAsnAsnIleVal 420
DB 1427 AACCAACAACAACAGTGTGATCATCCAGAGACTCAAAACCAACAACAATATATGTTGTG 1486
OY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB 1487 GATCATTCGATGATCAATTAATAAACAAGAACATGTTGACACAAAGGACAAATTAACAGCAGA 1546

```


OY 441 Asplys 442
 DB 1547 GNCANG 1552

RESULT 4
 AAA47753
 ID AAA47753 standard; cDNA; 1715 BP.

XX AAA47753;

DT 16-NOV-2000 (first entry)

DE VRN2 nucleic acid.

XX Vernalization gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 RW breeding; pollination; cultivation; ss.

OS Arabidopsis thaliana var 'Columbia'.

PN W0200044918-A1.

PD 03-AUG-2000.

PF 28-JAN-2000; 2000MO-GB00248.

PR 28-JAN-1999; 99GB-0001927.

PA (PLAN-) PLANT BIOSCIENCE LTD.

PI Dean C. Gendall A;

DR WPI: 2000-499333/44.
 P-PSDB: AAB00061.

PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 XX Claim 5; Page 71; 105pp; English.

CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.

SO Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;

Alignment Scores:

Pred. No.: 2,886-233 Length: 1715
 Score: 2285.00 Matches: 428
 Percent Similarity: 97.51% Conservative: 3
 Best Local Similarity: 96.83% Mismatches: 9
 Query Match: 96.09% Indels: 2
 DB: 21 Gaps: 1

US-09-890-220-2 (1-445) x AAA47753 (1-1715)

OY 1 MetCysArgGlnAsnGlySerProGluValIleSerThrSpglu 20
 DB 233 AGGTATGAGCAAAATGTCGCGCAATCTCACCAGAGAGTATTTCACTGATGAG 292
 OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 299 AATCTCTGATATATGTAACCTGTGACATAATAACATCTTCACCTTGGCTCTCA 352
 OY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 353 GGCAACCCATCGTTCTCTGCGAAGATGCTTGAACACTACAAAATTTGGGCAAGCGCAAAAGA 412

OY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnThrLeuGln 80
 DB 413 AAGTCAGATCTACTGGGATGTGATTTTCACTATAGAGATTGATATATCATTTACA 472
 OY 81 LysThrGluValArgLysPheCysSerCysProPheCysSerMetLeuGlySerPhe 100
 DB 473 AGAAGTGAAGTAGGAGGATGTCTGTGTCATTTTGTCTATAGCTATGTCGATGCTTC 532
 OY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheLysLeu 120
 DB 533 AAGGGCTGCATTTTCATTTGAAATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTT 592
 OY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 593 TTGGAAGATACCAACAGTTATATTTCTGTAACCTTAATCTTCATTTATTTAGGA 652
 OY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
 DB 653 GAAGAGATGATGATGATTAATTTGACCCCTCTCTCTGCTGCGAAACCTGTAAGGCT 712
 OY 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValLysPheLeuProLeuAsp 180
 DB 713 AGACAAAGAGGTGGCAGAAATACACACAGACACTTAAGATGCTTTTACCGTTGGAT 772
 OY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
 DB 773 TCACCCAGTTTACGTCATATGCGACAGAAATGGAATTCGCCCTGCTGATATGAAACCT 832
 OY 201 GlyLeuGlyTyrProGluValThrGluLeuValGlyGlnPheGluMetThrSerAsnIle 220
 DB 833 GGTTAGATATCCCGAGGCAACAGAGCTGTGTCGACATTTGAGATGCTAGCAACAT 892
 OY 221 ProProIleIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 DB 893 CCACCAAGCCATGCCACACTCTCTGAGCGTGGCTAAAGTATATTAACAACCGAA 952
 OY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGlyArgSerGluAlaArgSer 260
 DB 953 GCTGTGTCCTGCTACTAGACAAAGAAATATCTGCTGACGATGAGGCTGAGAAC 1012
 OY 261 HisLeuLeuGlnGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 DB 1013 CACCTACTCTTCAGAAAGCCAAATCTATCTATCTCACAGAGTCACCAATGGGCTT 1072
 OY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAlaAspAspValAlaAspPhe 300
 DB 1073 GAGCAAGTATCTGATCGAGTAGCGAGATGAAGTCGATGAGCATGTCACAGATTTT 1132
 OY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 DB 1133 GAAGATCGCCAGATGCTTATGATCTTTGTGATGTGAATGAATGAATGAATTCATG 1192
 OY 321 HisLeuTyrPAsnSerPheValArgLysGlnArgValIleAlaAspLysHisLeuSerTrp 340
 DB 1193 CATCTTTGGAATCGTTTGTATAGAAACAAAGGTTATGAGATGGTCATATCTCTGG 1252
 OY 341 AlaCysGluAlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPhe 360
 DB 1253 GCATGCAAGTATTTTCAAGATTTTACGAAAGAGTTTCACTGTTACATCACACTTTC 1312
 OY 361 TrpCysTrpArgLeuPheLeuLeuLysLeuTyrPAsnHisGlyLeuValAspSerAlaThr 380
 DB 1313 TGGTGTGGAGATTTGTTTGTATTAACATATGAAACATGAGCTTTCACATCAGCCACC 1372
 OY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
 DB 1373 ATCAACAATGCAATATCATCTCGAGATTCGCCGTAATAATCA-----GTCAATAC 1426
 OY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 DB 1427 AACACACACACAGTGTGATCATCCACTGACTCAAAACACACACACATATCAATGTGG 1486

OY 421 Asph1SPROASnAsp1LeAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
 DB 1487 GATCATCCGATGACATATAAACAAGAACAAATGTTGACACACAGACATATACACAGACA 1546
 OY 441 AspLys 442
 DB 1547 GACCAAG 1552

RESULT 5

AAAA47755
 ID AAA47755 standard; cDNA; 1737 BP.

AC AAA47755;

DT 16-NOV-2000 (first entry)

DE Aberrantly spliced VRN2 nucleic acid.

KX Vernalization gene: VRN2; plant characteristic: flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 breeding; pollination; cultivation; ss.

OS Arabidopsis thaliana var 'Columbia'

PN W0200044918-A1.

PD 03-AUG-2000.

PE 28-JAN-2000; 2000MO-GB00248.

PR 28-JAN-1999; 99GB-0001927.

RA (PLAN-) PLANT BIOSCIENCE LTD.

RI Dean C. Gendall A;

DR WPI: 2000-499333/44.

PS P-PSDB; AAB00062.

PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 PS
 PS Claim 8; Page 74-75; 105pp; English.

CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant. This cDNA was generated from an aberrantly spliced VRN2
 CC gene.

SO Sequence 1737 BP; 543 A; 330 C; 359 G; 505 T; 0 other;

Alignment Scores:

Pred. No.: 1,966-230 Length: 1737
 Score: 2258.50 Matches: 428
 Percent Similarity: 95.78% Conservative: 3
 Best Local Similarity: 95.11% Mismatches: 9
 Query Match: 94.97% Indels: 10
 DB: 21 Gaps: 2

US-09-890-220-2 (1-445) x AAA47755 (1-1737)

OY 1 MetCysArgGlnAsnGlySerValAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 233 ATGTGTAGGAGAGATGTCGCGGAATCTCACCGGAGAGATTTCAACGATGAG 292
 OY 21 AsnLeuLeuIleTyrCysLysProValAlaGluLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 293 AATCTCTGTATATATGTAAACCTGTGACATATATACATCTTTCACCTTCGCTCTCTA 352

OY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrIleGlyAlaLysArgLysArg 60
 DB 353 GGCAACCATCGTTTCTGCCAAGATGCTGTAGTACAAAATGTGGGCAAGCGCAAAAGA 412
 OY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnThrLeuGln 80
 DB 413 AAGTCAAGATCTACGCGGATGATGTTTCACTATAGAGATTTGATATATACATTACA 472
 OY 81 LysThrGluValAlaGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 DB 473 AGACTGAAAGTTAGGAGAGATTTGCTTGTGCATTTTGCCTATGCTATGCTGACTTC 532
 OY 101 Lys-----GlyLeuGlnPheHisLeuAsnSerSerHisAspLe 113
 DB 533 AAGTGGGCAACTATACACTGAGGCGTGCATTTTCATTGAAATTCATCTCATGATTT 592
 OY 113 upheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe 133
 DB 593 ATTTGAAATTTGGTTTCAGACTTTTGGAGAAATACACAGCATTAATGTTCTGTAAACT 652
 OY 133 uAsnSerPheIlePheGluGluGluGlySerAspAspLysPheGluProPheSerLe 153
 DB 653 TAATTCCTTCATATTTGAGAGAGAGAGATGATGATTAATTTGAGCCCTTCTCTCT 712
 OY 153 uCysSerLysProArgLysArgAlaGlnArgGlyLysArgAsnAsnThrArgArgLeuLys 173
 DB 713 CTGCTCGAAACCTCGTAAAGCGTAGCAAAAGATGGCGCAAAATMACCGACGAGACTTA 772
 OY 173 sValCysPheLeuProLeuAsnSerProSerLeuThrAsnGlyThrGluAsnGlyLLeuH 193
 DB 773 AGTATGCTTTTACCGTTGGATTCACCCAGCTTTAGCTATGGCACGAAATGGAATGTC 832
 OY 193 IleuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyL 213
 DB 833 CCTGCTGAATATGGAACCCGCTGTTTNGATATCCCGAGCGCAACAGACTTGCCTGGACA 892
 OY 213 nPheGluMetThrSerAsnIleProPheAlaIleAlaHisSerSerLeuAspAlaGlyL 233
 DB 893 ATTTGAGATGACTGACCAACATTCACCGACCATAGGCCACTTCTCTGAGCGCTGGTC 952
 OY 233 aLysValIleLeuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAl 253
 DB 953 TAAAGTTATTTAAACAACCGAGAGCTGTGCTCCTACTAAGACAAAGATATATCTGC 1012
 OY 253 AGluArgSerGluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHis 273
 DB 1013 TGAGCGATCAGAGCGCTGGAAGCCACTACTTCTTCAAGAAACCGCAATTTCTATCATTC 1072
 OY 273 sArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293
 DB 1073 CAGAGTCCAGCCAAATGGCGCTTGAGCAAGTATATGCTGATCGGAGATAGGAGATGAA 1132
 OY 293 LAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
 DB 1133 CGATGACGATTTGCAGATTTTGAAGATGCCAGATGCTTGTGATGACTTGTGCATGTCAA 1192
 OY 313 nLysAspGluLysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArgValAl 333
 DB 1193 TTAAGATTAAGAAAGATCATGCTTGTGAACTGCTTGTGAAAGAAACAAAGAGGTAT 1252
 OY 333 eAlaAspGlyHisIleSerThrPalacysgLualaPheSerArgPheTyrGluLysGluLe 353
 DB 1253 AGCAGATGCTATATTCCTTGGCGATGCTGAAGATTTTCAAAATTTTTCGAGAAAGACTT 1312
 OY 353 uHisArgTyrSerSerLeuPheThrPcysThrArgLeuPheLeuIleLysLeuThrPasnH 373
 DB 1313 GCACGTGTACTACATCACTCTTCTGCTGTGGAGATTTGATTGAATTAACATATGAAACCA 1372
 OY 373 sGlyLeuValAspSerAlaThrIleAsnAsnGlyAsnThrIleLeuGluAsnGlyArgAs 393
 DB 1373 TGAAGTGTGACTACAGCCACATCAACAACTGCAATACATCTCTCGAAGATTCGCGTAA 1432

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153718.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 9,16e-195
Score: 1923.00
Percent Similarity: 90.108

Length: 1497
Matches: 370
Conservative: 3

Best Local Similarity:	89.37%	Mismatches:	7
Query Match:	80.87%	Indels:	34
DB:	21	Gaps:	3
US-09-890-220-2 (1-445) x AAC51581 (1-1497)			
QY 43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer 62			
DB 322 CCATCGTTCTTCCAAAGTACGCTTGAACATAAATGGACCAAGCGCAAAAGAAAGTCA 381			
QY 63 ArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGlnTyrThr 82			
DB 382 AGATCTACTGGAGTGTGATTTTCACTATATAGGATTTGTAATACACATTACGAAACT 441			
QY 83 GluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPheLysGly 102			
DB 442 GAACTTAGGAGAGATGTTCTTCCATTTTGTCTATGCTATGTTGTTAGCTTCAAGGG 501			
QY 103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 122			
DB 502 CTGCATTTTCATTGCAATTCATTCATGATTTTATTGATTTGAGTTCAAG----- 552			
QY 123 GluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGlu 142			
DB 553 ---TACATCTGTGAATCCAAACCATATATGTTTGACG----- 588			
QY 143 SerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArgGln 162			
DB 589 -----TCGAACCTGTGAAGCGGAGACAA 612			
QY 163 ArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182			
DB 613 AGAGGTGGCAAAATTAACACCGAGAGACCTTAAGTATGCTTTTACCGTTGATTCACCC 672			
QY 183 SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArgGlyLeu 202			
DB 673 AGTTTAACTAATGGCACAGAAATGGAATCCCTCTTAAGATGGAACCCGTGTTTA 732			
QY 203 GlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPro 222			
DB 733 GGATATCCCGAGGACAGAACTTGGCGACATTTGAGATGACACGACAACTTCCACCA 792			
QY 223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVal 242			
DB 793 GCATATGCCACTCTTCTGTGACGCTGGCTGAAGTTATATGCAAGCAAGCACTGTG 852			
QY 243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259			
DB 853 GTCCCTGTACTAAGACAGAAAGTTATGCTGACGATCAGAGGCTAAGAGTACTGCC 912			
QY 260 -----SerHisLeuLeuGlnLysArgGlnPheTyrHis 271			
DB 913 ATATTAAACCGAGAAATTTCTTCCAGCACACTCTTCTTCAAGAAAGCCAAATCTATCAT 972			
QY 272 SerHisArgValGlnProMetAlaLeuGlnValMetSerAspArgAspSerGluAsp 291			
DB 973 TCTCACAGAGTCCAGCCAAATGCGCTTGACCAAGTATATCTGACCGGAGATAGCAGAT 1032			
QY 292 GluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAsp 311			
DB 1033 GAAGTCGATGACGATGTTGCGAGATTTGGAAGATCGGCAAGATCTTGATGACTTGTGAT 1092			
QY 312 ValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArg 331			
DB 1093 CTGATATAAGATGAAAAGCAATTCATGACATCTTTGGACACTGTTGTAAGAAACAAAGG 1152			
QY 332 ValIleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLys 351			
DB 1153 GTTATACAGATGCTCATATTTCTGCGCATGGAAGCATTTTCAAGATTTTACAGAAA 1212			
QY 352 GluLeuHisArgTyrSerSerLeuPheTyrPheTyrArgLeuPheLeuIleLysLeuTyr 371			
DB 1213 GAGTTGACCGCTTACTCATCTCTCTGTTGTGAGATTTGTTTGTGATTAACCTATAGG 1272			

OY 372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
 |||||||
 DB 1273 AACCATGGACTTGTGACTACACCATCAACAATGCAATACATCTCGAGAAATTGG 1332
 OY 392 ArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAspArgProSerAsp 411
 |||||||
 DB 1333 CGTATAGCTGACACACACACCAACAACAACAGTGTGATGTCCAGTGC 1392
 OY 412 SerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnAsnAsn 431
 |||||||
 DB 1393 TCAACACCAACAACAACAATAACATTTGTGATCATCCCAATGACATTAACAACAACAAT 1452
 OY 432 ValAspAsnLysAspAsnAsnSerArgAspLysValIleLys 445
 |||||||
 DB 1453 GTTGACAACAAGCAATTAACGACAGACAAAGTAATTA 1494

RESULT 7

AAA47752
 ID AAA47752 standard; DNA; 6338 BP.

XX AAA47752;

XX 16-NOV-2000 (first entry)

XX VRN2 nucleic acid.

XX Vernalization gene; VRN2; plant characteristic; flowering time;

XX leaf size; leaf shape; shade avoidance response; reproduction;

XX breeding; pollination; cultivation; ds.

XX Arabidopsis thaliana var 'Landsberg erecta'.

XX WO200044918-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000MO-GB00248.

XX 28-JAN-1999; 99GB-0001927.

XX (PLAN-) PLANT BIOSCIENCE LTD.

XX Dean C. Gendall A;

XX MPI; 2000-499333/44.

XX Isolated vernalization gene VRN2 is used to produce transgenic plants

XX with altered vernalization response, flowering time, leaf size and/or

XX shape or shade avoidance response for maximized reproductive success

XX Claim 9; Page 68-70; 105pp; English.

XX Isolated nucleic acid sequences obtained from the VRN2 locus of a

XX plant encode polypeptides which are capable of affecting one or more

XX vernalization responses such as, flowering time, leaf size and/or

XX shape or the shade avoidance response of a plant into which the

XX nucleic acid is introduced. Introducing such sequences into plants

XX to alter these characteristics maximises the reproductive success of

XX the plant.

XX Sequence 6338 BP; 1932 A; 1076 C; 1134 G; 2193 T; 3 other;

Alignment Scores:

Pred. No.: 3 76e-182 Length: 6338
 Score: 1813.50 Matches: 440
 Percent Similarity: 43.58% Conservative: 1
 Best Local Similarity: 43.48% Mismatches: 4
 Query Match: 76.26% Indels: 569
 DB: 21 Gaps: 13

US-09-890-220-2 (1-445) x AAA47752 (1-6338)

OY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluValIleSerThrAspGlu 20
 |||||||
 DB 2342 ATGTGTAGGCAAAATTTGTGCGCGGAATCTTCAACCGAGAGATGATTTCAACTGATGAG 2401
 OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 |||||||
 DB 2402 AATCTCTGATATATGTATGTAACCTGTGACATATATACATCTTCCACCTGCTCTCTA 2461
 OY 41 GlyAsn----- 42
 |||||||
 DB 2462 GGCACAGATGATTTGCTTCTCTCTCATCATATATAGCTCAGTAATCTTTCATCTCTG 2521
 OY 42 ----- 42
 DB 2522 TGTAGATCCACCACTAATAGTTGAGTTTGTAGCTGATATAGCTGATTCATGCGCA 2581
 OY 43 -----ProSerPhe 45
 |||||||
 DB 2582 GTGTGTGCTTCTTTTGTCTCTAAATTTGACCTGTGTGTGTGTCAGCCATCGTTT 2641
 OY 46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysSer----- 62
 |||||||
 DB 2642 CTTCCAGATGCTTGAACATCAAAATTTGAGCAAGCGCAAAAGAAAGTA-TGTTTCTT 2700
 OY 62 ----- 62
 DB 2701 CTTGATGATGAGCTGCTACAGTGAATGATATTTATCTACTTAATGCAAGCTGATG 2760
 OY 62 ----- 62
 DB 2761 ACCATATTTATCTTGTGTGATGATGATGACATATGATGATGTTCTTGTGTCAGT 2820
 OY 63 -----ArgSerThrGlyMetVal 68
 |||||||
 DB 2821 CTATAACTTACATTTTATATAAATGCTGTTTGTGATGATGATGATGATGATGATGAT 2880
 OY 69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 |||||||
 DB 2881 GTTTCACTATATAGATTTGATATACATATACAGAAACGAAAGGTAGCTTTTCT 2940
 OY 83 ----- 83
 DB 2941 GTTCGTGCAAAATTCGATGCAATGCTATGTTCTGATGATGTTGTTATTTACTA 3000
 OY 84 -----ValArgGlnLysPheCysSerCysProPheCysSerMetIle 96
 |||||||
 DB 3001 TTTTTCCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
 OY 96 uCysGlySerPheLys----- 101
 |||||||
 DB 3061 ATGTGTAGCTTCAAGGTGGGCACTATTAACAAGTGTCTTCCGGGCTTTCATA 3120
 OY 101 ----- 101
 DB 3121 TCTAACCTGTGAATGCTACTGCTGTTTCATGCTGTATATCTTACTGTTGTTACTAT 3180
 OY 101 ----- 101
 DB 3181 ATTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3240
 OY 102 -----GlyLeuG 104
 |||||||
 DB 3241 GAGAAACATATGCCAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
 OY 104 ImpPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu----- 120
 |||||||
 DB 3301 AATTCATTTGATTCATCATCATGATTTTGAATTTGAGTCAAGGATGTGTTTGA 3360
 OY 120 ----- 120
 DB 3361 TGGAAATTTCTGATTTGCTATGCTTATATATAGATGATGATGATGATGATGATGATGAT 3420
 OY 121 -----Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136

```

Db 3421 CTATTGTAGCTTTGGGAAAGATACACAGCTTAACTTCTGTAAACCTTAATTCCTTC 3480
Oy 137 ILepheglu----- 139
Db 3481 ATATTGAGGCACTTACTTTAACTGGTTAATGGAAATCCGATAGCTGCGAAAT 3540
Oy 140 -----glu1glu1yseraspaspaspasphegl 149
Db 3541 TTTGTTATATATCCATTCCTTATTTGATAGAGAAAGAGATGATGAGATTAATTTGA 3600
Oy 149 uPropheserleu----- 153
Db 3601 GCCCTTCTCTCTGTGTAACCCGACAGACCCCTTCGATTAAATACCTTAATAGCAGTAC 3660
Oy 154 -----Cyseryl 156
Db 3661 TCCCTTCTCTCTGTGTAACCTTCGTAATCCACCATTAATGTTTTGCGAGCTCGAA 3720
Oy 156 sProArgLysArgGlnArgGlyArgAsnAsnThrArgArgLysValCysPh 176
Db 3721 ACCTGTAAGCGGAGCAAGAGGTGCGAGAAATACACGAGACTTAAAGTATGCTT 3780
Oy 176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGlnAsnGlyIleThrLeuAs 196
Db 3781 TTTACCTGTGATTCACCCAGTTAACTAATGCGACAGAAATGGAATCACCCTTCTTA 3840
Oy 196 nAsp----- 197
Db 3841 TGATGTAATAATCATATCTTCTGTGCTTCTGTGCTTACACTTACATTAACAG 3900
Oy 197 ----- 197
Db 3901 AAGAAGATACAAAGCCCTGATGTTAGTTTTGTACTCTCTCCGATCTCTTGCGGA 3960
Oy 198 -----glyasnarglyle 202
Db 3961 GCGATATGTTACCGAAGATGATGACAAATTAATGCGATGCTACAGGAACCGGTGTT 4020
Oy 202 uGlyTyrProGlnAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIlePro 222
Db 4021 AGATATATCCGAGGCAACAGARCTCTGACAAATTTGAGATGACCAACATTCGCC 4080
Oy 222 alaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAla 242
Db 4081 AGCCATAGCCCTCTCTCTGAGCGCTGTCTAAAGTTATATGACAAAGCGTGT 4140
Oy 242 lValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259
Db 4141 GGTCCTGCTACTAAGCAAGAAAGTTATCTGTCGCGATCAGAGGCTAAGAGTTGCT 4200
Oy 259 ----- 259
Db 4201 TCATCATGACACCCCGTCAATTAATTAACATTCCTGTTTACAAATGTTCTTCATTT 4260
Oy 260 -----SerHisleu 263
Db 4261 ATGATATAGTATTATAGTACAGCCATATTAACCGAAATTTCTTCCAGCCACTACT 4320
Oy 263 uLeuGlnLysArgGlnPheTyrHisSerHisArgVal----- 275
Db 4321 TCTTCAGAAAGCCCAATTCATATCTTCACAGAGTCCAGGTGATCCAAATTCCTTGACC 4380
Oy 275 ----- 275
Db 4381 TACTCTTAGGCATTTTCTTAATGCTCATGATGATATCTTAACAAGCATACTTGT 4440
Oy 275 ----- 275
Db 4441 TTGTCTCATCAAAATTTGATTTGATGATGATGATATCAACGCAAAATAGTATATGCA 4500
Oy 276 -----Gln-ProMet 279

```

```

Db 4501 TGTGTCTCCGTTTATTTGCCACTAACCAAAAAATGATGTTTCTGTACAAACCAATGG 4560
Oy 279 lAleuGlnLysValMetSerAspArgAspSerGlnAspGluValAspAspValAla 299
Db 4561 CGCTTGAGAGAAATATGCTGACCGGGATGACGGAGATGAAGTCGATGACGATGTGCG 4620
Oy 299 sPheGluAspArg----- 303
Db 4621 ATTTGAAGATGCCAGGTATTCATGATTTCTTTCGCTCATTAATAATAGCAACAGA 4680
Oy 304 -----GlnMetleu 306
Db 4681 AATGGTATATGATGTAACCTGTAATGCGTTTGAACCTTAATAAAAAAGCTGCGATGCTT 4740
Oy 307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPhe 326
Db 4741 GATGACTTTGTGATGATGATTAAGATGAAGAAAGCAATTCATGATCTTTGGAACCTGTT 4800
Oy 327 ValArgLysGlnArgVal----- 332
Db 4801 GTAGAAAACAAAGGTA-ACATCTTCTTACACATGACAGACAGCAAAAAACCTTAG 4859
Oy 332 ----- 332
Db 4860 TCTTACATTCACATACCTGTCTAAATGATTTTGTCTTATGCACTTGACCTCAATTAATGAT 4919
Oy 333 -----IleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSer 346
Db 4920 TCTTATGTTTCAGGGTATATAGCAGATGTCATATTTCTTGCGCATGTGACATTTCA 4979
Oy 347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheThr----- 361
Db 4980 AGATTTACGAGAAAGAGTTGACCGCTTACTCATCTCTTCTGTAATATATAGTACACC 5039
Oy 361 ----- 361
Db 5040 AATCATATACAGACATTAACCTACATCAATCTGTTTCTGAAAAAAATA 5099
Oy 362 -----CysThrArgLeuPheLeuIleLysLeuThrAsnHisGlyLeuValAsp 377
Db 5100 AAAATTTCCAGTGTGAGATGTTTGTATTAACATTAATGAGACATGATGCTGAC 5159
Oy 378 SerAlaThrIleAsnAsnAsnSerValAspArgProSerAspSerAspThr 397
Db 5160 TCAGCCACATCAACAACTGCAATACCATCTCTGAGAAATTCGCGTAAATAGTCACACACC 5219
Oy 398 ThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAspThr 417
Db 5220 ACCACCAACCAACAAACAAACAGTGTGATGTCGCCAGTACTCAAAACCAACACAT 5279
Oy 418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnValAspAsnLysAspAsn 437
Db 5280 AACATTTGGGTCTATCCCAATGACATTAACAAAGAAAGTGTACACAGACAT 5339
Oy 438 AsnSerArgAspLysValIleLys 445
Db 5340 AACAGCAGAGCAAAAGTAATTA 5363

```

RESULT 8
 ID AAA47754 standard; DNA: 5895 BP.
 AC AAA47754;
 XX
 DT 16-NOV-2000 (first entry)
 DE
 XX VRN2 nucleic acid.
 XX
 KW Verbalization gene: VRN2; plant characteristic: flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ds.
 XX
 OS Arabidopsis thaliana var 'Columbia'

XX WO200044918-A1.
 XX 03-AUG-2000.
 XX 28-JAN-2000; 2000WO-GB00248.
 XX 28-JAN-1999; 99GB-0001927.
 XX (PLAN) PLANT BIOSCIENCE LTD.
 XX Dean C, Gendall A;
 XX WPI; 2000-499333/44.
 XX
 XX Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximizing reproductive success
 XX
 XX Claim 9; Page 72-74; 105pp; English.
 XX
 XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.
 XX
 XX Sequence 5895 BP; 1742 A; 1027 C; 1070 G; 2056 T; 0 other;

Alignment Scores:

Pred. No.: 9,986-166 Length: 5895
 Score: 1659.00 Matches: 424
 Percent Similarity: 42.25% Conservative: 4
 Best Local Similarity: 41.86% Mismatches: 12
 Query Match: 69.76% Indels: 577
 DB: 21 Gaps: 14

US-09-890-220-2 (1-445) x AAA47754 (1-5895)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 1905 ATGTCTAGGCAAAATTTGCGCGCAATCTCCACCGAGAGATGATTTCACACTGATGAG 1964
 QY 21 AsnLeuLeuIleTyCysLysProValArgLeuTyAsnIlePheHisLeuArgSerLeu 40
 DB 1965 AATCTCTTGATATATGTGAACCTGTGCACTATATACATCTTCCACCTGCTCTCTA 2024
 QY 41 GlyAsn----- 42
 DB 2025 GGCAATCGATGATTTGGCCTTCCTCTCATCATTTTAGCTTAATCTTTCATCTCC 2083
 QY 42 ----- 42
 DB 2084 TGTGTAGATACCAACCACTAATAGTTGAGTTGCTAAGCTGATTATGCTGACTGATGC 2143
 QY 43 -----Pro 43
 DB 2144 GAGTGTGTCTCTTTGTCTCTCTAATGTTATTGAACTGTGTGTGTGTGTGAGGCCA 2203
 QY 44 SerPheLeuProArgLysAsnTyLysIleGlyAlaLysArgLysArgLysSer--- 62
 DB 2204 TCGTTTTCGCCAAGATGCTTGACTACAAATAATGGGCAAGCGCAAAAGAAAGTA-TGC 2262
 QY 62 ----- 62
 DB 2263 GTTCTCTCTGAATGATGTTGCCACAGTATATGTTATTATCTTACTCTAATATGAA 2322
 QY 62 ----- 62
 DB 2323 GCTGATGAACATATTATCTTGTGTGATGATATGACATAATGAATGCTTCTCTTTG 2382

QY 63 -----ArgSerThrGly 66
 DB 2383 TTTCATGCTATACACTTATATTTTACAAAATTTGTTTCTGCTAGCGTCAAGATCTACTGGG 2442
 QY 67 MetValValPheAsnTyLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 DB 2443 ATGTAGATTTCACACTAATAGATGTATATATACATTACAAAGAACTGAAGTTAGTCT 2502
 QY 83 ----- 83
 DB 2503 TTTTCTGTCTTCGACAAAATTCGATGTCATGCTATGTTCTCTAGATGATTTGTTAT 2562
 QY 84 -----ValArgGlnAspCysSerCysProPheCysSe 94
 DB 2563 TTACTATTTTTCCTGTAATGTCAGCAGCACTTAGGAGAGATTCTGTGCTATTTGCTC 2622
 QY 94 rMetLeuCysGlySerPheLys----- 101
 DB 2623 TATGCTATGCTGACTTCACAGTGGGCAACTATTACAACTGAGGTTCTTCCGGGGCCT 2682
 QY 101 ----- 101
 DB 2683 TTCAATATCTACACTGTGAATGCTACTGCGCTTAATGCTATATACCTTTCACGTGTGG 2742
 QY 101 ----- 101
 DB 2743 TTACATATTTTGTGTTTCTTGTCTCTCTTCTGCTCTTTTAACTGCTGAGTGTGCTC 2802
 QY 102 -----G 102
 DB 2803 TTTATCTGAGAAAACATGTTCCAGTTCGAGCTTACATCCATGCTCTGTGTATGACGAG 2862
 QY 102 lyleGlnPheHisLeuAsnSerHisAspLeuPheGlnPheLysLeuPhe- 121
 DB 2863 GGCTCAATTTATTTGAAATTCATCATGATTTATTTGAATTTAGTTCAAGATATG-T 2921
 QY 121 ----- 121
 DB 2922 GGTTTTATGAAATTTCTTGTGTTTGGCTATGCCGTTAGTAATGAGCTTATAGTTAAAG 2981
 QY 122 -----GluGluTyGlnThrValAsnValSerValLysLeuA 134
 DB 2982 GGTCTTCTCTATTTGAGCTTTTGGAAATATACAGACAGTATATGTTCTCTGTAAGACTTA 3041
 QY 134 snSerPheIlePheGlu----- 139
 DB 3042 ATTCCTTCATATTTGAGGTCAGTACTTTAAACTGTGTTAATGGGAAATCTATAGCTG 3101
 QY 140 -----GluGlnGlySerAspAsp 146
 DB 3102 GTCAAAATTTGCTTATATTCATCCTTATTTGTAAGTGAAGAAAGAAAGTATGATGAT 3161
 QY 147 LysPheGluProPheSerLeuCys----- 154
 DB 3162 AATTTGAGCCCTCTCTCTCTG-CTAATCTCAGAACCCCTTGATTAATACCTTAATA 3220
 QY 154 ----- 154
 DB 3221 GCAGTAATCTCTGCTTTCTTCTGTCAGTACTTCTCTAATAATCAACCAATGTTTTC 3280
 QY 155 -----SerLysProArgLysArgLysArgGlnArgLysAsnThrArgArgLeuLys 173
 DB 3281 AGCTGAACACCTCGTAAGGCTGACAAAGAGGTGGCAGAAATTAACCCGAGACTTAA 3340
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGlnAsnGlyIleThr 193
 DB 3341 GTATGCTTTTACCGTGTGATTCACCACTTACCTAATAGCAGAAATGAAATTTGCC 3400
 QY 194 LeuLeuAsnAsp----- 197
 DB 3401 CTGCTGAAGATGTAATAATCAATCTTCTGTGTGATTCGTTGTGCTTGAACCTTC 3460
 QY 197 ----- 197

D	b		3461	ATTTTACAGAAAGAATACAATGACATGCCGATGTGGTATTGGTTTGACTCTCCCGCATTC	3520
O	y		198	-----GlyAs	199
D	b		3521	TTCCTGTGAGGTAATGTTACCAGAACGTGATGACAAAATTAAATGCATGCCTACAGAAA	3580
O	y		199		
D	b		3581	CCTGGTTTAGGATATCCCGAGCAACAGAGCCTTGCTGGACAAATTTGGATGATGATCAACA	3640
O	y		219	nlllepprolalilealhsserterleuasapalagyalalyvalilleautlrse	239
D	b		3641	CATTCACACGAGCATAGCCCACTCTCTCGACGCTGGCTTAAAGTTAATATAACAC	3700
O	y		239	rglualValAlproAlarhlrlystrrralgylseuseralaglunasergluaAlAr	259
D	b		3701	CGAAGTGtGhrccttgctaTAAGACAAAGAAATTATCTGTGAGCATCAGAGCTAG	3760
O	y		259	g-----	259
D	b		3761	AAGTTGTTCATCATGACACCCCGTCATCATTAATACCATACCTGTTGTACAAATGTT	3820
O	y		260	-----Se	260
D	b		3821	CTTCCATTATGATGAATGTTTACGTGCTAGTGCATTAATACCGAAAAATTTCTCCAG	3880
O	y		260	rHlslleuleuleuglnLysArgGlnPhetYrHlslserHlsArgValGln	276
D	b		3881	CCACCTACTTCTTCAGAAAAAGCCCAATCTCATATCTTCACAGAGTCCAGGGATCCAACT	3940
O	y		276	-----	276
D	b		3941	TCCTCACCTACTCTTATAGCATTTCTTAAATGTCATGATGATTCATTATCAAAGC	4000
O	y		276	-----	276
D	b		4001	ATACTGGTTTGTTCTCATCTAAATTTGTATTTTGATTCTGTATGTATCAACGCAAAAA	4060
O	y		276	-----	276
D	b		4061	ATTATGTCATGTGTCTCCGTTTTATATGCCCACCTAACCAAAAACCTGATTTCTTGTA	4120
O	y		277	---PrometalaleuGlInValMetSerAspArgSerpSerGluaspGluValaspas	295
D	b		4121	CMACCCAATGGCCCTTGAGCAGATGATGTCTGATCGGAGTAAAGGAGATGAAGTCAAGTA	4180
O	y		295	pAspValAlaAspPheGluAspArg	303
D	b		4181	CGATGTGGCAGATTTTGAAGATCGCCAGATTCATCATGATTTCTTCTCGTTCATTAA	4240
O	y		303	-----	303
D	b		4241	TAGGCAACAGAAAAATGATATACGATGTAACTTGCTAATGGCTTTGAAAATTAAAAAGC	4300
O	y		304	--GlmelLeuAspAspPheValIAspValIasnLysaspGluLysglnPheMetHlsLeut	323
D	b		4301	TGCAGATGCTTGATGATGACTTTGGTGGATGTAATAAAGATGAAGAACCAATCATGCACTTT	4360
O	y		323	rPaenSerPheValIargLysgln	330
D	b		4361	GGAACCTGTTGTGAAGAAAAACAAGGTAACACTCTCTTACACTTGAACACACACAAAA	4420
O	y		330	-----	330
D	b		4421	AGACCTTATGCTTACATTCATACCTGTCTAAATGATTCGTTATAGCACTTGAGCT	4480
O	y		331	-----ArgValIleAlaaspGluHlsIleSerTrpalacysglu	343
D	b		4481	CAAAATTATGATTTGATGTTGACAGGGTTATAGAGAGTGCATATCTCTTGGGAGTGA	4540
O	y		344	AlaPheSerArgPheYrGluLysglnLeuHlsArgYrIserSerleuPheTrp	361

Dd		4541	GATATTTCAGACTTTTACGAAAGAAGTGCTACTGTACTCATCACTCCTTCTGGTAATAT	4600
Oy		361	-----	361
Dd		4601	AAGTACCACCAACATATACAGACATATACACTATCATTTTGTTCGTTTCGA	4660
Oy		362	-----CSTPACGleupheleullelylsleuTPASnHs61	374
Dd		4661	AAGAAAATAAATAAATTCACAGTGTGGAGATTTGTTTTGATTAACAATAGAACACATGG	4720
Oy		374	yleuValASPserAlatThrIleAsnAnscYsaNthrIleLeuGIuaScYsaTgaSnSe	394
Dd		4721	ACTGTGCGACTGACGCCACCATCCAAACACTGCATATACCATCTCTGAGAAATTTGCCGTAAATAC	4780
Oy		394	rSerAspThrThrThrThrAsnaAsnaSseValaspArgProSerAspSerAntH	414
Dd		4781	CTCA-----GTCACTAACCAACCAACACACAGTGTGATCATCCACAGTACTCANAC	4834
Oy		414	fAsnaSnaSnaSnIleValaspHisProAsnAspIleasnaSnlysaSnaSnaValaspAs	434
Dd		4835	CAACAAACAATTAACATTGTGCGATCATCCGATGACATATAAAAAACAGAACAAATGTTGACAA	4894
Oy		434	nLysaspAsnaSnaSsePargaspLys	442
Dd		4895	CAAGGACAATTAACGACGAGCACAG	4919
RESULT_9				
AAZ36947				
ID	AAZ36947	standard;	cdNA:	2280 BP.
XX	AAZ36947:			
XX	13-MAR-2000	(first entry)		
DE	cdNA encoding a Mpcl protein having flowering regulating activity.			
XX				
KW	Flowering regulating activity; Mpcl; flowering; germination;			
KV	super early flowering mutation; altered flowering time;			
KW	flowering regulating gene; food crop; vegetable; flowering inhibition;			
KW	productivity; ss.			
OS	Arabidopsis thaliana.			
PH	Key	Location/Qualifiers		
FT	CDS	310..2145		
FT		/*tag= a		
FT		/product= "flowering regulating protein"		
PN	EP967278-A2.			
PD				
XX	29-DEC-1999.			
PF	28-JUN-1999;	99EP-0305077.		
PR	26-JUN-1998;	98JP-0180065.		
PR	24-JUN-1999;	99JP-0179043.		
PA	(MITSUI CHEM INC.)			
PI	Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T.			
DR	WPI: 2000-064612/06.			
PS	P-PsDB: AAI53932.			
PT	Novel DNA used to produce transgenic plants with altered floral			
PT	regulation which can have increased crop yields -			
PS	Claim 6; Page 16-21; 53pp; English.			
CC	The present sequence encodes a protein having a flowering regulating			
CC	activity, which is designated MPCL. The genomic sequence is given			
CC	in AAZ36948. A rice Mpcl is also disclosed in the specification. The			
CC	rice and Arabidopsis cDNAs show significant homology with each other.			

CC A naturally occurring mutation of the Mpc1 gene eliminates normal
 CC flowering regulating ability of plants, and leads to flowering
 CC immediately after germination (super early flowering mutation). The
 CC Mpc1 polynucleotide sequence can be used to produce plants with altered
 CC flowering times in comparison with wild type plants, by enhancing
 CC or inhibiting the expression of the flowering regulating gene. Antisense
 CC polynucleotides can be used to reproduce the effects of the mutated Mpc1
 CC gene. This alteration can be used to increase the yield of food crops.
 CC Flowering inhibition of vegetables increases their productivity.

xx Sequence 2280 BP; 663 A; 445 C; 521 G; 651 T; 0 other;

Alignment Scores:

Pred. No.:	6,95e-88	Length:	2280
Score:	923.00	Matches:	216
Percent Similarity:	43.23%	Conservative:	52
Best Local Similarity:	34.84%	Mismatches:	108
Query Match:	38.81%	Indels:	244
DB:	21	Gaps:	9

US-09-890-220-2 (1-445) x AA636947 (1-2280)

```

QY 1 MetCysArgGlnAsnGlyArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
   ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 310 ATGTCCATGGAAGCTCCGCTGCTGCTATTCGGAAAGAGAGAGATTCCTGTAAGAG 369
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 370 AGCTTGCGCTGCTATTCGTAAGGCTTTGAACTCTCAATATTCATTCACAGCGCTGCTAT 429
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 AGGAATCCCTGTTCTTCAGGAGATTTGCATATTAAGATGAGCAAAACATATAAAG 489
QY 60 ----- 60
Db 490 AGAATACAAATGACTGTATTCCTCTCGGGCGCTATAGATGCTGGGTACAAATCAAAA 549
QY 60 ----- 60
Db 550 TTATTCCTCTGTATATTTTGTGGCAAGACTCGTTTCTCTAAGCTGTGCTGATAT 609
QY 60 ----- 60
Db 610 TCTGCAGTATATAGTTGACGTGAGCATGATATCTAAGTGTGATGGGGTTGATGA 669
QY 60 ----- 60
Db 670 GTTACTCAAGCCCAAGCAACTTCTCTCCCTGATATGATAGATGTCATTTGAGGCA 729
QY 60 ----- 60
Db 730 AAATGAGATCATCTGCTATCTTGTATTACGCTTTGCTGGTGGCCAAAATTTCAATTT 789
QY 60 ----- 60
Db 790 GGCATTGATTCAGGCAAGATTCATTCAGGAATATAGAGAGCATTTGTTATGAGCAAA 849
QY 60 ----- 60
Db 850 ATACCTTGCATCACTGATATGCTGTGCGAGAAATACACAAACATGAGACTTGGAGAG 909
QY 61 ----- 61
Db 910 AGAGTAGACACAGTCTCTCTTGTGAATGACAGCTTGCCTTCATTAAGCTAAAGTCAATG 969
QY 64 Ser ----- 64
Db 970 AGTAGGAAGAAAGTGTCTCGATTTCAGTTCAGGCAACATTCACATCCGAGCTCTGCG 1029
QY 64 ----- 64
Db 1030 CAGCAAGTAGCAAGTCCACATATCTGACAGAGAAGTTGGTCAACGAAAAATCTCTTAT 1089

```

```

QY 64 ----- 64
Db 1090 ACTTATTTTCATTAATGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1149
QY 65 ThrGlyMetValIlePheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGluVal 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1150 ACAGAAATATGATTTTCACTACAGATACATATACAAACAAATTCAGAGACTGAGTA 1209
QY 85 ArgGlyAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGlyLeuGln 104
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1210 ACTGAAGACTTTCTTCTGCAATCTGCTTGAATATGCGCAGTTTCAGGCTGAGA 1269
QY 105 PheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluTyr 124
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1270 TATCTCTGCTCCATCAACCCAGCATCTCTCAATTCGAGTTTGGGTACAGAAATTT 1329
QY 125 GlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGlnLysSerAsp 144
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1330 CAGCGCGTAATGTCTCCCTCAAGACTGAGACATGATATCCAGTTAATGAGGATGAC 1389
QY 145 ---AspArgLysPheGluProPheSerLeuCysSerLysPro---ArgLysArgArgGln 162
   ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1390 GTTGACCCCAACACCAACTTCTTTTCTTCTCCAAAATTCAGACGGAGGCA 1449
QY 163 ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1450 AAGACTCAGTACGAGCTCAAGGCA-----GGGCT 1482
QY 183 SerLeuThrAsnGly-----ThrGluAsnGlyIleThrLeuLeuAsn 196
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1483 CATCTTGATTAAGTTGGAGGTGCTAGATAGACATGATGATGATGATGATGATGATGAT 1542
QY 197 AspGlyAsnArg-----GlyLeuGlyTyrProGluAlaThrGluLeuAlaGln 213
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1543 GAGAGAGCCGATACCACTGGAAGAGCATTC-----GAAAGATTTGGGGT 1590
QY 214 PheGluMetThrSerAsnIlePro-----ProAlaIleAlaHisSerSerLeu 229
   ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1591 GCTGAGTCTGCGCAAGAGTCTCTCTGACAGCATCTCTGACAGCTGCAATCATGTGG 1650
QY 230 AspAlaGlyAlaLysValIleLeuThrSerGluAlaValAlaProAlaThrLysThrArg 249
   ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1651 GATCCAGATTAATGCTACGATACGATGGAATGCAATGTTGCAAAAACGAG 1710
QY 250 LysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArgGlnPhe 269
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1711 AAAATATCTATAGAACGGTGGCATTTGAGGAACCGAAGCCTCTCAGAGAGACAGTTC 1770
QY 270 TyrHisSerHisArgValGlnProMetAlaLeuGlnGlnValMetSerAspArgAspSer 289
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1771 TTCACCTCTCATGAGCTCAGCCCATGGCTCTAGAACATGATCTTGGACCGGATAGT 1830
QY 290 GluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPhe 309
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1831 GAAAGATGAAGTTGATGATGATGTCGACATTTGAAGATAGAAAGATGCTCATGATTTTC 1890
QY 310 ValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPheValArgLys 329
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1891 GTTGATGTGACTAAAGATGAGAAACAGATGACATGTCGATGTCGTTGTGAGAG 1950
QY 330 GlnArgValIleAlaAspGlyHisIleSerThrLysGluAlaPheSerArgPheThr 349
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1951 CAGCGAGTATTAAGCATGCTGCTCATTCATGGCATGCGAGGCACTTCTCAAGATTTGAC 2010
QY 350 GlnLysGluLeuHisArgTyrSerSerLeuPheThrCysTyrArgLeuPheLeuLys 369
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2011 GGACCCATCATGCTTCAACACCGCATTTGATTTGCTGAGAGAGTATTAAGTGA 2070
QY 370 LeuThrAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGln 389
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2071 CTGTGAAACCAAGGCTTCTTGATGCTCCGAAACATGAACAACTGTAATACCTTCTCGAA 2130

```

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rng

Page 14

```

RESULT 10
AAZ36953
AAZ36953 standard; cDNA; 2248 BP.
XX AC AAZ36953;
XX DT 13-MAR-2000 (first entry)
XX DE cDNA encoding an Os-MPC1 protein having flowering regulating activity.
XX KW Flowering regulating activity; MPC1; flowering; germination;
XX KW super early flowering mutation; altered flowering time;
XX KW flowering regulating gene; food crop; vegetable; flowering inhibition;
XX KW productivity; ss.
XX OS Oryza sativa.
XX FH Key Location/Qualifiers
XX FT CDS 86..1900
XX FT /*tag= a
XX FT /product= "flowering regulating protein"
XX PN EP67278-A2.
XX PD 29-DEC-1999.
XX PF 28-JUN-1999; 99EP-0305077.
XX PR 26-JUN-1998; 98JP-0180065.
XX PR 24-JUN-1999; 99JP-0179043.
XX PA (MITA ) MITSUI CHEM INC.
XX PI Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;
XX DR WP1: 2000-064612/06.
XX DR P-PSDB: AAY53933.
XX PR Novel DNA used to produce transgenic plants with altered floral
XX PR regulation which can have increased crop yields .
XX PS Claim 6; Page 40-44; 53pp; English.
XX CS The present sequence encodes a protein having a flowering regulating
XX CC activity, which is designated Os-MPC1. An Arabidopsis MPC1 is also
XX CC disclosed in the specification. The rice and Arabidopsis cDNAs show
XX CC significant homology with each other. A naturally occurring mutation
XX CC of the MPC1 gene eliminates normal flowering regulating ability of
XX CC plants, and leads to flowering immediately after germination (super
XX CC early flowering mutation). The MPC1 polynucleotide sequence can be
XX CC used to produce plants with altered flowering times in comparison
XX CC with wild type plants by enhancing or inhibiting the expression of
XX CC the flowering regulating gene. Antisense polynucleotides can be used
XX CC to reproduce the effects of the mutated MPC1 gene. This alteration
XX CC of vegetables increases their productivity.
XX SO Sequence 2248 BP; 641 A; 454 C; 493 G; 660 T; 0 other;

Alignment Scores:
Score: 3 91e-86 Length: 2248
Prod. NO.: 906.50 Matches: 209
Percent Similarity: 42.688 Conservative: 59
Best Local Similarity: 33.288 Mismatches: 109
Query Match: 38.128 Indels: 251
DB: 21 Gaps: 7

US-09-890-220-2 (1-445) x AAZ36953 (1-2248)
0Y 1 MetCysArgGlnAsnCysArgAlaIysSerSerProGluGluValIleSerThrAspGlu 20
Db 86 ATGTGGCGCCACACGACGAGGCTGCGGTCTCTCCGATGAGACACACTGTCAGCTGAGGAA 145

```

OY	2	AsnLeuLeuIleTyCysLysProValArgLeuValPasnIlePheHisLeuArgSerLeu	46
Dd	146	AGCTTGCAATTATACGTGCAAGCGGCAGTGTATAATATCATTCAGCCGCATCATAT	205
OY	41	GlyAsnProSerPheLeuProArgCysLeuAsnTrpLysIleGlyAlaLysArgLysArg	60
Dd	206	AAAATHCCTGCTTTCTTCAAAGAATGCCCTCTTACAGATTCACGAAGCGAAGAAC	265
OY	61	LysSerArgSerThr-----	65
Dd	266	AGGAGCGCTGATACCATATACACTTCTTGAGGACAAATAAAGACTGGCGGCACAAT	325
OY	65	-----	65
Dd	326	ATCTTTCCTCTTANGTCTGTAGCTAGACCCTACATAATATGTTCACCTTGAAGGCAT	385
OY	65	-----	65
Dd	386	TCTCGCATATATGATTACATTCAGTCGCTGTTGTGTGACTTCCTTTCATGAAATTTGAAAT	445
OY	65	-----	65
Dd	446	AAGACTAGACTGAAGCAACATCATCTGCATTCGTGATGGAAGAACTTAGCAACCTCCGA	505
OY	65	-----	65
Dd	506	GCTTGCAGCCTTAATATTTATTCCTWTATTCAGCTGTGGAGAGCTGAGCAAACTTTTGATGAC	565
OY	65	-----	65
Dd	566	AATAACTGTTTCGGGAACCATGTGGAAAGCTCTACTCCAAAAAGTTGAAGGAAGTGT	625
OY	65	-----	65
Dd	626	TTCGGGGTAAAAATACCAATGCAATCTTCTGCTTCATCTTTGGSAATTTGTGACTTA	685
OY	65	-----	65
Dd	686	AGTTTGGGACATACCGTGGAAATGTCCTCCACGGTTGAGATGACCCCAAGCTTCTTAGAG	745
OY	65	-----	65
Dd	746	CCAAATTTCTGGAGAGTACAGTGTCTGACATTTTGTCTCAGAAAGTTGATGCTACT	805
OY	65	-----	65
Dd	806	GGTTCAATTCACACTGCAGATTAGCATATCTGCTCAGAGCGTGGTGCAAAAAGACATGTCC	865
OY	65	-----	65
Dd	866	GAGTCCTTATAGTGTATTATCANATATAGTGTGCCACCTTCGTGTCATTTGACACATATT	925
OY	66	-----GlyMetValValPheAsnTrpLysAspCysAsnAsnTrpLeuGln	80
Dd	926	ATAAGGTGAGATCTGGCAATGTGCTTTTAACTACAAATACATCAATATATCTATGACAA	985
OY	81	LysTrpGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe	100
Dd	986	AAAACCGAAGTACAGAAATTTTTCTTGCCCAATTTTGTGGTACCATGTGGCACCTTT	1045
OY	101	LysGluLeuGluPheHisLeuAsnSerSerHisAspleuPheGluPheGluPheLysLeu	120
Dd	1046	AAGGCTCTAGATGTCACCTTAAAGCAGCATCGCATACCTTTTCCATATATAGATTGGATA	1105
OY	121	PheGluGluTrpClnTrpValAsnValSerValLysLeuAsnSerPheIlePheGlu---	139
Dd	1106	TCGAAAGATGCCAGCTGTATTAATGTATGTCTGAGACCTGATCTTGGGAACAGAGCTT	1165
OY	140	--GluGluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArg	158
Dd	1166	TTGGCTGAGGGAGGTT--GATCCAAAGACATCCAATTTTGTACCGCTCAAGATTAAAG	1222
OY	159	LysArg--ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu	177

```

Db      1223  AAGCTAAAGGTTGGAATCTCAAGTATAAATTAGGACATGATCAATCATATTTGTG 1282
QY      178   ProleuAspSerPro---SerleuThrAsnGlyThrGluAsnGlyIleThrleuAsn 196
Db      1283  GATTTCAGATCACTGGAAGATGCCAGGACGATCTGGAAGATTACGTCAGAGGAA 1342
QY      197   AspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMet 216
Db      1343  AATGTTAGT----- 1351
QY      217   ThrSerAsnIleProProAlaIleAlaHisSerSerleuAspAlaGlyAlaVal--- 235
Db      1352  -----TCTGTACACACGCTCTGTGATCTGTGATCTGTCAATTCATTACAC 1393
QY      236  -----IleleuThrSerGluAlaValAlaProAlaThrIleGlyThrArgLysleuSerAla 253
Db      1394  GGTACCAATCTTTCCGACCAACACTGTTACAGTTGGGAACACAAAGAGCTGCTGT 1453
QY      254  GluArgSerGluAlaArgSerHisleuLeuGlnLysArgGlnPheTyrHisSerHis 273
Db      1454  GAACGAGCTGATCCCAAGATCGCAGCTCCCTACAAAAACGCCAGTTCTTTCATCTCAC 1513
QY      274  ArgValGlnPrometAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluVal 293
Db      1514  AGGCTCAACCAATGCGATGAGAGCAAGCTTTCTCAGATCGTGAATGAGATGAGT 1573
QY      294  AspAspArgValAlaAspPheGluAspArgGlnMetleuAspAspPheValAspValAsn 313
Db      1574  GATGATGACATTCGATTTTGGAGATGAGAAATGATGATTTGTTGTCATCTTACA 1633
QY      314  LysAspGluLysGlnPheMetHisleuThrAsnSerSerPheValArgLysGlnArgValIle 333
Db      1634  AAAGCGAGAAACTATTATGATATGTGGAATTTATTTGTTCCGAAACAAAAGGCTACTA 1693
QY      334  AlaAspGlyHisIleSerTerPalAcysGluAlaPheSerArgPheTyrHisArgLysGluLeu 353
Db      1694  GCGGATGGCCATATTCCTGGGCGATGCGAAGCATTCGCAAGTTCATGAGCAACAACAT 1753
QY      354  HisArgTyrSerSerleuPheTyrPcysTrrPargLeuPheLeuIleLysleuThrAsnHis 373
Db      1754  GTACAAATCCAGCTCTACTATGATGTTGGAGTTTATGTTAACTCTGGAAACAC 1813
QY      374  GlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393
Db      1814  ACTCTACTGATGATGCCGACCATGATGATGCTTCGACACAAATCTTGAAGGCTACTGAAAC 1873
QY      394  SerSerAspThrThrThrThrAsn 401
Db      1874  GGAAGCTCGATCCAAAGAAAAT 1897

RESULT 11
AAA47758 standard; cDNA; 558 BP.
XX      AAA47758;
AC      AAA47758;
XX      16-NOV-2000 (first entry)
DE      At Hyp 2245035 (ATFCA7_4) (modified) cDNA.
XX      Vernalization gene; VRN2; plant characteristic; flowering time;
XX      leaf size; leaf shape; shade avoidance response; reproduction;
XX      breeding; pollination; cultivation; ss.
OS      Arabidopsis thaliana.
PN      WO200044918-A1.
XX      03-AUG-2000.
XX      28-JAN-2000; 2000WO-GB00248.
XX

```

```

PR      28-JAN-1999; 99GB-0001927.
XX
PA      (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI      Dean C, Gendall A;
XX
DR      WPI; 2000-499333/44.
XX
P-PSDB; AAB00065.
XX
PT      Isolated vernalization gene VRN2 is used to produce transgenic plants
PT      with altered vernalization response, flowering time, leaf size and/or
PT      shape or shade avoidance response for maximized reproductive success
XX
PS      Disclosure; Page 76; 105pp; English.
XX
CC      Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC      plant encode polypeptides which are capable of affecting one or more
CC      vernalization responses such as, flowering time, leaf size and/or
CC      shape or the shade avoidance response of a plant into which the
CC      nucleic acid is introduced. Introducing such sequences into plants
CC      to alter these characteristics maximises the reproductive success of
CC      the plant. This cDNA represents another Arabidopsis thaliana sequence
CC      having homology to the VRN2 gene.
XX
SQ      Sequence 558 BP; 191 A; 122 C; 112 G; 133 T; 0 other;

Alignment Scores:
Pred. No.: 1,55e-38 Length: 558
Score: 451.50 Matches: 99
Percent Similarity: 67.18% Conservative: 32
Best Local Similarity: 50.77% Mismatches: 43
Query Match: 18.99% Indels: 21
DB: Gaps: 7

US-09-890-220-2 (1-445) x AAA47758 (1-558)
QY      237   LeuThrSerGluAlaValAlaProAlaThrIleGlyThrArgLysleuSerAlaGluArgSer 256
Db      16   CTGACACAGAAAGCTAAGTCCCT-----GCTAAGGATCA 51
QY      257   GluAlaArgSerHisleuLeu---LeuGlnLysArgGlnPheTyrHisSerHisArgVal 275
Db      52   AAGGCTACAGACCATCTACTCTGCTTCATAAACCCAGTCTTATCATCTCCGCAACCGGT 111
QY      276   GlnPrometAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAsp--- 294
Db      112   CAGCCATGTCCTTGACCAAGTATGTCTGACCGAGATGCGAAATGACCTGCAACAA 171
QY      295   --AspAspValAlaAspPheGluAspArgGlnMetleuAspAspPheValAspValAsn 313
Db      172   AATGATGATGCTGACATCTCGAAGAACCCGATGCTTAATGCTTCATGATGATGAGAT 231
QY      314   Lys---AspGluLysGlnPheMetHisleuThrAsnSerSerPheValArgLysGlnArgVal 332
Db      232   GAAATCGTAGCAGACAGATTCATTAACCTTTGGAACCTCTTGTAAACAGCAAGAGAT 291
QY      333   IleAlaAspGlyHisIleSerTerPalAcysGluAlaPheSerArgPheTyrGlyLysGlu 352
Db      292   GTTGCAGATGCTCATATTCCTTGGCATGTGTAAGCATTTCTCAAGATTACACCTGCAAGAG 351
QY      353   LeuHisArgTyrSerSerleuPheTyrPcysTrrPargLeuPheLeuIleLysleuThrAsn 372
Db      352   CTGGCCAGTAATCTTATCTACTGCACTTCCTGAGAACATTCATGATGATCAAAATAGGAT 411
QY      373   HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392
Db      412   TATGACCTCTTGCAGACAGTCAACATGAAATGCAATATCATCATCATCATCATCAT 465
QY      393   AsnSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSer 412
Db      466   AATATCTCAACTACACAGATGACATTAACATTAACACACACA-----AGG 510
QY      413   AsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsn 427

```


XX EP67278-A2.
 PN
 XX 29-DEC-1999.
 PD
 XX 28-JUN-1999; 99EP-0305077.
 PF
 XX 26-JUN-1998; 98JP-0180065.
 PR 24-JUN-1999; 99JP-0179043.
 XX
 PA (MITA) MITSUI CHEM INC.
 XX
 XX Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;
 PI WPI, 2000-064612/06.
 DR P-PSDB; AAY53932.
 XX
 XX Novel DNA used to produce transgenic plants with altered floral
 PT regulation which can have increased crop yields -
 XX
 PS Example 1: Page 26-34; 53pp; English.

CC The present sequence encodes a protein having a flowering regulating
 CC activity, which is designated Mpc1. The cDNA sequence is given
 CC in AA236947. A rice Mpc1 is also disclosed in the specification. The
 CC rice and Arabidopsis cDNAs show significant homology with each other.
 CC A naturally occurring mutation of the Mpc1 gene eliminates normal
 CC flowering regulating ability of plants, and leads to flowering
 CC immediately after germination (super early flowering mutation). The
 CC Mpc1 polynucleotide sequence can be used to produce plants with altered
 CC flowering times in comparison with wild type plants, by enhancing
 CC or inhibiting the expression of the flowering regulating gene. Antisense
 CC polynucleotides can be used to reproduce the effects of the mutated Mpc1
 CC gene. This alteration can be used to increase the yield of food crops.
 CC flowering inhibition of vegetables increases their productivity.

XX Sequence 5580 BP; 1496 A; 999 C; 1075 G; 2010 T; 0 other;

Alignment Scores:
 Pred. No.: 2.4e-32 Length: 5580
 Score: 407.50 Matches: 187
 Percent Similarity: 24.79% Conservative: 47
 Best Local Similarity: 19.81% Mismatches: 87
 Query Match: 17.14% Indels: 626
 DB: 21 Gaps: 18

US-09-890-220-2 (1-445) x AA236948 (1-5580)

QY 65 ThrglyMetValIalPheAsnTyrlYsAsPcySaSnAsnThrlLeuGlnlyThrgluVal 84
 DB 2722 ACAGGAATGTAGTTCCTTAACATACAGATCTATACAAACAATTCGACAAAGCTAAGG- 2780
 QY 84 ----- 84
 DB 2781 TAACAGTATATTATTTAACTGTTTCAATCCCAATGTCTATATTTCATCCGTTACCC 2840
 QY 85 ----- ArgGluAspCysSerCysPro 91
 DB 2841 TAACCTGTACGCTATATGTTGCTATGTCTTCGACGATCACTGAAGACTTTTCTGTCCA 2900
 QY 92 PheCysSerMetLeuCysGlySerPheLys----- 101
 DB 2901 TTTCGCTTAGTAATAATGTGCAGATTCAAGGTGACCTTCATTTCCATTCATTCATCC 2960
 QY 101 ----- 101
 DB 2961 TCTTAGTCAAAGATACACCTGTAGTACTGTTTGTAGTATGCAATCTTTCTTTT 3020
 QY 102 ----- GlyLeuGlnPheHsiLeuAsnSerHisAspLeuPheGlu 115
 DB 3021 CTCCTCAATCATGTGTAGGCGCTGAGATATCATTTGCCATCAACCCAGATCTCTCAAT 3080
 QY 116 PheGlu-PheLysIeu----- 120

DB 3081 TTTCAGTTTGGGTTGTAGCTTTAAATTCAGTTAACTGTTTATCTTTTATT 3140
 QY 121 ----- PheGluIuTyrgInThrValAs 128
 DB 3141 TTGTGGGTCCCACTAATCTGCTTACTTGTAGTACTGACAAATTCAGGCGTAAA 3200
 QY 128 nValSerVallys----- 132
 DB 3201 TGTCCTCCCAAGACTGAGACAATGATATCCAAAGGTTAGAAACATCTGTGTTGCATTT 3260
 QY 132 ----- 132
 DB 3261 ATGTTTCACTTCTGCTGCTATATCTATAGCTGTACAAATTCATTTTCATTTA 3320
 QY 133 ----- LeuAsnSerPheIlePheGluIuGlySer----- 143
 DB 3321 AACTAATATCCTCCATGAGGTGTTGACTTTGTGTGTTAATAAGGAACTGCAATCTT 3380
 QY 144 ----- AspAspAspIly 147
 DB 3381 TAGTTGCTATTTGTCACACTATGATCTTGCTATTTGCTTAATAGCTGATGAGAAATTA 3440
 QY 147 sPheGlu----- 149
 DB 3441 ACT-CAAAATGACATCGCTGTTCTGTTACTTTTGTGGCATGAGACCGCAAAAGCTCG 3499
 QY 149 ----- 149
 DB 3500 ACTGTAGATAAAGTCTCGATTAATATAGAGTCAAACTAATTAAGTAGTTGTTTC 3559
 QY 149 ----- 149
 DB 3560 TACAAATATTTCTATGCTTTGTTACTTTTCCATTTATGATTACTTACACACATTT 3619
 QY 149 ----- 149
 DB 3620 TCTAATGTTAATGTTTATTAATAAATCTGCTCAGGTTAATGAGATGACGTGACCCA 3679
 QY 150 ----- ProPhe----- 151
 DB 3680 AGCAGCAAACTTTCTTTTTCGTAAATATCTGCGCTATATGTTGCTTTTATATCTT 3739
 QY 152 ----- SerLeuCysSerIysProArgLysArgAr 161
 DB 3740 TCCACACTCTGTGTGACGCCATTAATAATTTCTCAATATGTGACAGTTCCAAAATAATTGAG 3799
 QY 161 gGlnArgGlyIuArgAsnAsnThrArgArgLeuLysVal----- 174
 DB 3800 ACGGAGGAGGCAAAAGAGTCAAGTACGAGGTCAGAGGCAAGGCGCTCATCTTGATTAGG 3859
 QY 175 -CysPheLeu-----ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyI 192
 DB 3860 TTTCGAGGTGCTAGATTAAGACTGARGGTATGTGTACTGAAATGACAGATTAATTTGAT 3919
 QY 192 eThrLeuLeu----- 195
 DB 3920 TTGTAGTATTTGGCTTTTGTGATGACAGACGCTGCTAGTTGATATTATTAACGATAT 3979
 QY 196 ----- AsnAspGlyAsnArgGlyLeuGlyTyPr 205
 DB 3980 TTACTTTGTTATGTCATATTTTGCATGCAACAACGTTGATGATTCATTTGACACAGCTTTTA 4039
 QY 205 oGluAlaIthrGluLeuAlaGlyGlnPheGluMet--ThrSerAsnIleProProAlaIle 224
 DB 4040 TTCTTACTTTCAGATGCTCATCTCTGTAGAGTGAGAAAGCGCAATCCACCTGGA-- 4096
 QY 225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValPro 244
 DB 4097 AAGCTTTCGAAAGAAATTTGGGGTCTGAG-----TCTGTCAAAAGATCTCTCCT 4147
 QY 245 AlaThr----- 246

Db 4148 GGCACGAGTCCTGCAGACGTCGATCATCATGTTGGGATCCAGATTAATGTGACGTCGATGCT 4207
 QY 247 -----LysThrArgLysLeuSerAlaGluArgSerCysLeu 258
 Db 4208 GGAAGTACAAATGTGACAGTTTGCAGAAAAGAGAAATATATATGAAAGCGTGCAGACTTG 4267
 QY 259 ArgSerHisLeu----- 262
 Db 4268 AGCAAGTATGTTGACTTCCTTTGCTGCTATCCCTCTTCTCAATTATTAATTAACATA 4327
 QY 262 ----- 262
 Db 4328 CATATGTCATGATGAGAAAATGTGCTTAAGTTTAACAAGTACCTGTTAATCCC 4387
 QY 263 -----LeuLeuGlnLys 266
 Db 4388 AATGATGTGAGTGAAGTTTTCAAAATTTTTCCTCCACGCGAGCCCTCCCTTCAGAG 4447
 QY 267 ArgGlnPheTyrHisSerHisArgVal----- 275
 Db 4448 AGACAGTCTCTCCACTCATCGACGTCAGTGAATCTTTTCTTACGCTCTGCTT 4507
 QY 276 -----GlnProMetAlaLeuGlu 281
 Db 4508 TGAAGATTCGAATGATTTTTCATTGCTATGCTACTATGCTGAGCCCATGGCTGAGAA 4567
 QY 282 GlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPheGlu 301
 Db 4568 CAAGTACTTTCGACCGGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4627
 QY 302 AspArg----- 303
 Db 4628 GATAGAGAGGCTATGTTTGAATTAATTTTCAACCGCATCAGTCTGGTGAATAA 4687
 QY 303 ----- 303
 Db 4688 ACCTCAGTGTGGTGTGATATATGTTTCATGTGAAGGAGAAAGAAATTTGAAGACTGG 4747
 QY 303 ----- 303
 Db 4748 GCATGGCGCAACGTTAGAGCAATTTGTAAATGTTGACAGATCAATAGAAAATATCTGA 4807
 QY 303 ----- 303
 Db 4808 GCAAGCCTCAGGTTGATATGAGACAGTAGAACAGATCATAGTCTTATATACACT 4867
 QY 303 ----- 303
 Db 4868 CATTAAGACGAGAGTGTCCGTTTGTACTGATCTAACATAGTGTGATTCATACATAG 4927
 QY 304 -----GlnMetLeu 306
 Db 4928 TTGTCTGATCTCCATATAGTAATAACGTTATTTCTATTAATTTCTTTCAGATGCT 4987
 QY 306 ValAspAspPheValAspValAspValAspGlnLysGlnPheMetHisLeuTrpAsnSerP 326
 Db 4988 CGATGATTTCTGTTGATGTGACTAAAGATGAGAAACATGATGACATGTGCAACGCTT 5047
 QY 326 eValArgLysGlnArgVal----- 332
 Db 5048 TGTAGAGAGAGAGGGGTA-TGTCTTATCTCTTTTCATGATGTCAGCTGAGATTTTCCA 5106
 QY 332 ----- 332
 Db 5107 GTATTAACATTTAGAGTGGCGCATGTAAGGTTTGATTAATTCCTGCTGGCTTCTCT 5166
 QY 332 ----- 332
 Db 5167 GGTAAAAAATAAACTGAACAATAGATTAACATGACATGCTCTCTGACTCAT 5226
 QY 333 -----IleAlaAspGlnHisI 338
 Db 5227 ATTAACGATTAACCTTGACAGTGTGTTGGACCTTTGACAGATTAATGACAGATGCTCAT 5286

QY 338 eSerTrpAlaCysGlnAlaPheSerArgPheTyrGlnLysGlnLeuHisArgTyrSer 358
 Db 5287 TCATATGGCATGACGAGGATTTCTCAATTTGCACGGACCATCATGATGTTGAAACCGCA 5346
 QY 358 TrpPheTrp----- 361
 Db 5347 CTTCATTTGGTAATTAACACTCTCATTTCTTCATTTGTTTTCAGTGTATCGAGAGA 5406
 QY 362 -----CysTrpArgLeuPheLeuI 368
 Db 5407 AAGCGTTTGTGATATAAAGTACGCTTTTGTGTAGTGTGAGAGAGTGTATGCT 5466
 QY 368 eLysLeuTrpAsnHisGlnLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLe 388
 Db 5467 GAACCTGTGAAACCGAGCTCTTCTGATGCCGAGAACCATGAACTGTAATACCTTCT 5526
 QY 388 uGlu 389
 Db 5527 CGAA 5530
 Db 5527 CGAA 5530
 RESULT 13
 AAA47757
 ID AAA47757 standard; cDNA; 325 BP.
 AC AAA47757;
 DT 16-NOV-2000 (first entry)
 DE A1163743 Expressed sequence tag (modified) cDNA.
 KW Vernalization gene; VRN2; plant characteristic; flowering time; poplar;
 KW leaf size; leaf shape; shade avoidance response; reproduction; EST;
 KW breeding; pollination; cultivation; expressed sequence tag; ss.
 OS Populus tremula.
 PN WO20044918-A1.
 PD 03-AUG-2000.
 PF 28-JAN-2000; 2000WO-GB00248.
 PR 28-JAN-1999; 99GB-0001927.
 PA (PLAN-) PLANT BIOSCIENCE LTD.
 PI Dean C. Gendall A;
 DR WPI; 2000-499333/44.
 DR P-PSDB; AAB00064.
 DR XX
 PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 PS Disclosure; Page 76; 105pp; English.
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant, encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximizes the reproductive success of
 CC the plant. This cDNA represents expressed sequence tag clone A1163743
 CC from poplar trees (Populus tremula x Populus tremuloides) which has
 CC homology in the C-terminal region of the Arabidopsis thaliana VRN2 gene.
 SO Sequence 325 BP; 102 A; 60 C; 76 G; 87 T; 0 other;
 Alignment Scores: 4,84e-30 Length: 325
 Pred. No.: 368.50 Matches: 76
 Score:

PR 28-MAR-1999; 99US-013149.
PR 30-APR-1999; 99US-0132048.

20-JUL-1999;	99US-0144814
21-JUL-1999;	99US-0144814
21-JUL-1999;	99US-0145086

```

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0148975.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149326.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```

```

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1,29e-26 length: 640
Score: 340.50 matches: 67
Percent Similarity: 80.00% conservative: 1
Best Local Similarity: 78.82% mismatches: 16
Query Match: 14,328 indels: 16
DB: 21 gaps: 2

US-09-890-220-2 (1-445) x AAC44147 (1-640)
OY 362 CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIle 381
Db 305 TCGTGGCGG-----TGG-----TGC 319
OY 382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
Db 320 AACACATGCAATACCATCTCGAGAAATTCGCCGTAAATGCTCAGACACCCACCCACCAAC 379
OY 402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleValAsp 421
Db 380 AACACAAACACAGTCGTGATCGTCCAGTCAACACCAACCAACAAATTAATCTGTGCAT 439
OY 422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
Db 440 CATCCCAATGACATTAACAACAACAAAGAACAAATGTTGACACAAAGACAAATTAACGACAGAA 499
OY 442 -LysValIleLys 445
Db 500 CAAAGTAATTA 512

RESULT 15
AAC51582
ID AAC51582 standard; DNA: 319 BP.
XX
AC AAC51582;
XX
DI 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69042.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

```


PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0125264.
PR 29-MAR-1999; 99US-0125785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152633.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158366.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Prod. No.:	1.25e-25	Length:	319
Score:	327.00	Matches:	61
Percent Similarity:	96.838	Conservative:	0
Best Local Similarity:	96.838	Mismatches:	2
Query Match:	13.758	Indels:	0
DB:	21	Gaps:	0

US-09-890-220-2 (1-445) x AAC51582 (1-319)

QY 383 AsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsn 402
DB 3 AACTCGAATACCATCCTCGAGAAATTGGCGTAATAGCTCAGACACCACCACCAAC 62
QY 403 AsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleValAspHis 422
DB 63 AACACACAGTGTGATGCTCCAGTCACCAACACCAACCAATTAACATTGTGATCAT 122
QY 423 ProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAspLys 442
DB 123 CCCAATGACATTAACACACAGAACATGTTGACCAACACGACATTAACAGCAGACAA 182
QY 443 ValIleLys 445
DB 183 GTAAATTAA 191

Search completed: June 19, 2003, 23:51:41
Job time : 377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:39:38 ; Search time 70 Seconds
(without alignments)
1949.588 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRONRAKSSPEEYISTDE.....INNNKVVNDKNNRDYIK 445

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144852_16502/app_query.fasta.1.583
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09890220.ecgn.1.140.etruncat_13062003_144852_16502 -NCPD=6 -ICPD=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Issued Patents_NA:*
2: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/BACKFILES1.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120.5	5.1	198	PCT-US95-10668-3	Sequence 3, Appl1
2	120.5	5.1	198	PCT-US95-10668-4	Sequence 4, Appl1
3	114	4.8	198	PCT-US95-10668-1	Sequence 1, Appl1
4	114	4.8	198	PCT-US95-10668-2	Sequence 2, Appl1
5	109	4.6	2237	US-08-914-999-7	Sequence 7, Appl1
6	107	4.5	1086	US-08-415-751-27	Sequence 27, Appl1
7	107	4.5	1086	US-08-415-751-28	Sequence 28, Appl1
8	107	4.5	5163	US-08-700-651-1	Sequence 1, Appl1
9	107	4.5	5163	US-08-928-361B-4	Sequence 4, Appl1
10	107	4.5	5318	US-08-700-651-2	Sequence 2, Appl1
11	107	4.5	5318	US-08-928-361B-3	Sequence 3, Appl1
12	107	4.5	5511	US-08-928-361B-2	Sequence 2, Appl1

13	107	4.5	7334	3	US-08-928-361B-1	Sequence 1, Appl1
14	105.5	4.4	1430	1	US-08-276-452A-25	Sequence 25, Appl1
15	105.5	4.4	1430	2	US-08-798-744-25	Sequence 25, Appl1
16	104.5	4.4	1680	2	US-08-700-152A-3	Sequence 3, Appl1
17	103	4.3	4423	4	US-09-541-782-5	Sequence 5, Appl1
18	103	4.3	4423	4	US-09-723-820-5	Sequence 5, Appl1
19	98.5	4.1	4757	4	US-09-572-191-1	Sequence 1, Appl1
20	98.5	4.1	4757	4	US-09-723-262-1	Sequence 1, Appl1
21	98.5	4.1	4757	4	US-09-723-219-1	Sequence 1, Appl1
22	98.5	4.1	7425	4	US-09-453-702B-212	Sequence 212, Appl1
23	98	4.1	1086	1	US-08-415-751-47	Sequence 47, Appl1
24	97	4.1	1690	1	US-08-276-452A-24	Sequence 24, Appl1
25	97	4.1	1690	2	US-08-798-744-24	Sequence 24, Appl1
26	96.5	4.1	2093	1	US-08-287-001A-1	Sequence 1, Appl1
27	96.5	4.1	2093	1	PCT-US95-09941-1	Sequence 1, Appl1
28	96	4.0	1956	4	US-08-559-896B-1	Sequence 1, Appl1
29	94	4.0	8920	2	US-08-446-855A-1	Sequence 1, Appl1
30	94	4.0	8920	4	US-09-150-741-1	Sequence 1, Appl1
31	93	3.9	700	4	US-09-236-097-11	Sequence 11, Appl1
32	93	3.9	5720	4	US-09-442-100-1	Sequence 1, Appl1
33	91.5	3.8	2136	4	US-09-177-248-1	Sequence 1, Appl1
34	91.5	3.8	2315	4	US-09-061-769A-1	Sequence 1, Appl1
35	91.5	3.8	2322	4	US-09-061-769A-4	Sequence 4, Appl1
36	91.5	3.8	2327	4	US-09-061-769A-3	Sequence 3, Appl1
37	91.5	3.8	5361	4	US-08-973-462-2	Sequence 2, Appl1
38	91.5	3.8	6152	4	US-08-973-462-1	Sequence 1, Appl1
39	91	3.8	1631	3	US-09-118-319-1	Sequence 1, Appl1
40	91	3.8	2288	2	US-08-568-459A-5	Sequence 5, Appl1
41	91	3.8	2288	2	US-08-487-826B-5	Sequence 5, Appl1
42	91	3.8	2288	4	US-09-210-288-5	Sequence 5, Appl1
43	91	3.8	3800	4	US-08-965-762-13	Sequence 13, Appl1
44	91	3.8	3800	4	US-09-911-927-13	Sequence 13, Appl1
45	91	3.8	3800	4	US-09-911-927-15	Sequence 15, Appl1

ALIGNMENTS

RESULT 1
PCT-US95-10668-3
Sequence 3, Application PC/TUS9510668
GENERAL INFORMATION:
APPLICANT: James Eberwine
TITLE OF INVENTION: A Method of Sequencing Proteins by
TITLE OF INVENTION: Epitope Ordering and Protein
TITLE OF INVENTION: Restriction Mapping
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10668
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/294,133
FILING DATE: August 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 198
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: No
PCT-US95-10668-3

Alignment Scores:
Pred. No.: 4,91e-06 Length: 198
Score: 120.50 Matches: 24
Percent Similarity: 65.08% Conservative: 17
Best Local Similarity: 38.10% Mismatches: 21
Query Match: 5.07% Indels: 1
DB: Gaps: 1

US-09-890-220-2 (1-445) x PCT-US95-10668-3 (1-198)

QY 382 AsnaScySasnrHlleuGluaScYsArGaSnSerSerAspThrThrThraSn 401
DB 5 AACCAACAACAACAACAACAACAACAAGATTCCAAACAACAACAACAACAAC 64
QY 402 AsnaSnSerValaSpArgProSerAspSerAsnThraSnAsnaSnlleValaSp 421
DB 65 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 121
QY 422 HisProAsnAspIleAsnAsnLysAsnAsnValaSpAsnLysAsnAsnSerArgAsp 441
DB 122 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 181
QY 442 LysValaIle 444
DB 182 AAGCTGATG 190

RESULT 2
PCT-US95-10668-4
Sequence 4, Application PC/TUS9510668
GENERAL INFORMATION:
APPLICANT: James Eberwine
TITLE OF INVENTION: A Method of Sequencing Proteins by
TITLE OF INVENTION: Epitope Ordering and Protein
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10668
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/294,133
FILING DATE: August 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 198
TYPE: Nucleic Acid

STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: No
PCT-US95-10668-4

Alignment Scores:
Pred. No.: 4,91e-06 Length: 198
Score: 120.50 Matches: 24
Percent Similarity: 65.08% Conservative: 17
Best Local Similarity: 38.10% Mismatches: 21
Query Match: 5.07% Indels: 1
DB: Gaps: 1

US-09-890-220-2 (1-445) x PCT-US95-10668-4 (1-198)

QY 382 AsnaScySasnrHlleuGluaScYsArGaSnSerSerAspThrThrThraSn 401
DB 5 AACCAACAACAACAACAACAACAACAAGATTCCAAACAACAACAACAACAAC 64
QY 402 AsnaSnSerValaSpArgProSerAspSerAsnThraSnAsnaSnlleValaSp 421
DB 65 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 121
QY 422 HisProAsnAspIleAsnAsnLysAsnAsnValaSpAsnLysAsnAsnSerArgAsp 441
DB 122 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 181
QY 442 LysValaIle 444
DB 182 AAGCTGATG 190

RESULT 3
PCT-US95-10668-1
Sequence 1, Application PC/TUS9510668
GENERAL INFORMATION:
APPLICANT: James Eberwine
TITLE OF INVENTION: A Method of Sequencing Proteins by
TITLE OF INVENTION: Epitope Ordering and Protein
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10668
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/294,133
FILING DATE: August 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 198
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: No


```

OY 261 HistLeuLeuGlnLysArgGlnPheYrHisSerHisArgValGlnProMetAlaLeu 280
DB 620 CATTAAACCTCCAAAAAGATTGAAATTTTCACAAACAGC-GTT-----TTA 666
OY 281 GlnGlnValMetSerArgAspSerGluAspGluVal-----AspAspAspValAla 298
DB 667 GAGTTTGTGTA-----GATACACATCATCATGATTAAATTTGTGAGACGACCAATGTGA 720
OY 299 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAspGlnArgValIleAla 317
DB 721 GAAGGACAAATATGAAAGTATATATATATATATATATATATATATATATATATATATAT 780
OY 318 -----GlnPheMetHisLeuTrpAsnSerPheValArgGlnArgValIleAla 334
DB 781 ACACACAAATCATTTCTCTCTCATTTCCATATGAAACATTCAAATCATCAATTTATGATTATA 840
OY 335 Asp-----GlnHisIleSerTrpAlaCysGlnAlaPheSerArgPheTyrGlu 350
DB 841 GATATTCAGGTGGTGGT-----GATCAGCTATATACAGAC 873
OY 351 LysGlnLeuHisArgTyrSerSerLeuPheTyrGluPheTyrGluPheLeuIleLysLeu 370
DB 874 CCACAAATTCATACCTATG-----TGGTGTGGT-TTGGTATTTGGTAAAT-TTG 921
OY 371 TrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu--- 389
DB 922 GGTCAAAAAGGTTTGAAAAGTTTATGACTCTCAATGATGAAATGCAATTTGCCAATAT 981
OY 390 ---AsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnAsnSerValAspArg 408
DB 982 TTAATTTTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1041
OY 409 ProSer-----AspSerAsnThrAsnAsnAsnAsnIle 419
DB 1042 CCAGATTAATTTTCCCTGATACATCTGAAAGATATATATATATATATATATATATATAT 1098
OY 420 ValAspHisProAsnAspIleAsnAsnValAspAsnValAspAsnValAspAsnSer 439
DB 1099 ---AATATATATATATATATATATATATATATATATATATATATATATATATATATAC 1155
OY 440 ArgAspLysValIleLys 445
DB 1156 AATAGTAGTATTTCAAAA 1173

RESULT 6
US-08-415-751-27
Sequence 27, Application US/08415751
Patent No. 5643772
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PHILLIPS, MOORE, LEMPJO & FINLEY
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: California
COUNTRY: United States of America
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 KB storage
COMPUTER: PC
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,751
FILING DATE: 03-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/071,880
FILING DATE: June 1, 1993
APPLICATION NUMBER: 07/891,301
FILING DATE: May 29, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Hana Dolezalova
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-1677
TELEFAX: (415) 324-1678
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cryptosporidium parvum

US-08-415-751-27

Alignment Scores:
Pred. No.: 0.00383 Length: 1086
Score: 107.00 Matches: 23
Percent Similarity: 58.33% Conservative: 12
Best Local Similarity: 38.33% Mismatches: 25
Query Match: 4.50% Gaps: 0

US-09-890-220-2 (1-445) x US-08-415-751-27 (1-1086)

OY 382 AanaAncysAnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
DB 345 AACACACCAACCAACTTATTCATCATTTACACACACCAACCAACCAACCAACCAAC 404
OY 402 AanaAncysAnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 421
DB 405 AACACACCAACCAAGGTCCAGGTAGCCACCAATAGCCACCAACCAACCAACCAACCA 464
OY 422 HisProAsnAspIleAsnAsnValAspAsnValAspAsnValAspAsnValAsp 441
DB 465 AATAGTTTCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACG 524

RESULT 7
US-08-415-751-28/C
Sequence 28, Application US/08415751
Patent No. 5643772
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PHILLIPS, MOORE, LEMPJO & FINLEY
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: California
COUNTRY: United States of America
ZIP: 94306-1840
COMPUTER READABLE FORM:

```



```

1 ZIP: 80303
2 COMPUTER READABLE FORM:
3 MEDIUM TYPE: Floppy disk
4 COMPUTER: IBM PC compatible
5 OPERATING SYSTEM: PC-DOS/MS-DOS
6 SOFTWARE: Patentin Release #1.0, Version #1.25
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: US/08/276,452A
9 FILING DATE: 18-JUL-1994
10 CLASSIFICATION: 435
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Caruthers, Jennie M.
13 REGISTRATION NUMBER: 34,464
14 REFERENCE/DOCKET NUMBER: 27-91A
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (303)499-8080
17 TELEFAX: (303)499-8089
18
19 INFORMATION FOR SEQ ID NO: 25:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 1430 base pairs
22 TYPE: nucleic acid
23 STRANDEDNESS: single
24 TOPOLOGY: linear
25 MOLECULE TYPE: cDNA
26
27 FEATURE:
28 NAME/KEY: CDS
29 LOCATION: 2..1312
30
31 FEATURE:
32 NAME/KEY: misc_feature
33 LOCATION: 41..112
34 OTHER INFORMATION: /note= "Derived amino acid sequence
35 OTHER INFORMATION: corresponding to the peptide sequence by protein
36 OTHER INFORMATION: microsequencing"
37
38 FEATURE:
39 NAME/KEY: misc_feature
40 LOCATION: 41..112
41 OTHER INFORMATION: /note= "Derived amino acids 14-24,
42 OTHER INFORMATION: 28, 30, 32-37 are identical to sequences obtained
43 OTHER INFORMATION: by protein microsequencing"
44
45 FEATURE:
46 NAME/KEY: misc_feature
47 LOCATION: 25..31
48 OTHER INFORMATION: /note= "Amino acids 25, 27, 29, and
49 OTHER INFORMATION: 31 are hydroxylated proline residues; amino acid
50 OTHER INFORMATION: 26 can be T instead of A."
51
52 US-08-276-452A-25
53
54 Alignment Scores:
55 Pred. NO.: 0.00927 Length: 1430
56 Score: 105.50 Matches: 28
57 Percent Similarity: 42.11% Conservative: 12
58 Best Local Similarity: 29.47% Mismatches: 32
59 Query Match: 4.44% Indels: 23
60 Db: 1 Gaps: 2
61
62 US-09-890-220-2 (1-445) x US-08-276-452A-25 (1-1430)
63
64 QY 372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
65 |||:||||| :|||: ||||| |||
66 Db 551 AACAAAGGCTACTCGGAGAAATTACAAACAACAACAACAACAAATGCTACTCGGAGAAATTAC 610
67
68 QY 392 ArgAspSerSerAsp:
69 |||:|||||:
70 Db 611 AACAAACAACAACAATGCTACTCCAAAGATTACAAACAACAACAATGCTACTCCAAAAAATC 670
71
72 QY 397 -----ThrThrThrThrAsnAsnAsnSerValAspArgProSer 410
73 Db 671 AACAAATATGTTACTCCGAGAAATTACATGAACAACAACAACGCGTCTCCGAGAGTTAC 730
74
75 QY 411 AspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsn 430
76 |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77 Db 731 AACAGCAACAACAACAACAACAATATTTTCTCCGAGAAATTACAAACAACAACAATATATACAAAT 790

```

```

OY      431 AsnVal-----AspSlnYsAspSnsAsnSerTrtGspLysVal 443
Db      791 AANGTGTTTCTCGAGATTATACACACAAATATATACATATATGTT 835

RESULT 15
US-08-798-744-25
Sequence 25, Application US/08798744
Patient No. 5830747
GENERAL INFORMATION:
APPLICANT: Chen, Chao-Guang
APPLICANT: Mau, Shiao-Lim
APPLICANT: Du, He
APPLICANT: Gane, Allison M
APPLICANT: Bagic, Anthony E
TITLE OF INVENTION: Plant Arabidogalactan Protein (AGP) Genes
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
City: Boulder
STATE: Colorado
COUNTRY: United States of America
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/798,744
APPLICATION NUMBER: 08/276,452
FILING DATE: 18-JUL-1994
PRIORITY INFORMATION:
PRIORITY APPLICATION DATA:
CLASSIFICATION: 435
APPLICATION NUMBER: 08/276,452
FILING DATE: 18-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Cartuberts, Jennie M.
REGISTRATION NUMBER: 34,464
REFERENCE/DOCKET NUMBER: 27-91A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303)499-8080
TELEFAX: (303)499-8089
TELEX: 49617824
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1430 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..1312
FEATURE:
NAME/KEY: misc.feature
LOCATION: 41..112
OTHER INFORMATION: /note= "Derived amino acid sequence
corresponding to the peptide sequence by protein
OTHER INFORMATION: microsequencing"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 41..112
OTHER INFORMATION: /note= "Derived amino acids 14-24,
31 are hydroxylated proline residues; amino acid
OTHER INFORMATION: 28, 30, 32-37 are identical to sequences obtained
by protein microsequencing"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 25..31
OTHER INFORMATION: /note= "Amino acids 25, 27, 29, and
31 are hydroxylated proline residues; amino acid
OTHER INFORMATION: 26 can be T instead of A"

```

US-08-798-744-25

Alignment Scores:

Pred. No.:	0.00927	Length:	1430
Score:	105.50	Matches:	28
Percent Similarity:	42.11%	Conservative:	12
Best Local Similarity:	29.47%	Mismatches:	32
Query Match:	4.44%	Indels:	23
DB:	2	Gaps:	2

US-09-890-220-2 (1-445) x US-08-798-744-25 (1-1430)

```
QY 372 AsnHISgLYLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
    |||:||||| :|||: ||||| ||| |||||
DB 551 AACATGGGCTACTCGAGATTACACACACACACACACATGGCTACTCGAGATTAC 610
QY 392 ArgAsnSerSerAsp----- 396
    |||:|||||:
DB 611 AACACACACACACATGGCTACTCGAGATTACACACACACATGGCTACTCGAAAAAATC 670
QY 397 -----ThrThrThrAsnAsnAsnAsnSerValAspArgProSer 410
    |||:|||||:
DB 671 AACATTAATGGTTACTCCAGATTACATGAACACACACACGCTTCTCCGAGAGTTAC 730
QY 411 AspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsn 430
    :|||:||||| ||||| ||||| ||| |||
DB 731 AACACCAACACACACACACATATTCTTCGAGATTACACACACATATAATACAAAT 790
QY 431 AsnVal-----AspAsnLysAspAsnAsnSerArgAspLysVal 443
    |||:||||| :|||: ||||| ||| |||
DB 791 AATGTTTCTCCGAGATTACACACACATATAATACATATATGTT 835
```

Search completed: June 20, 2003, 01:25:08
Job time : 95 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 20, 2003, 00:55:28 ; Search time 231 Seconds
(without alignments)
2826.862 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRACKSSPEEVISTDE.....INKNKNVDKDNNSRDYIK 445

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh
-MODEL=frame+ p2n model
-O=/cgn2.1/USFto_spool/US09880220/runat_13062003.144854.16586/app.query.fasta.1.583
-DB=Published_Applications_NA -OFM=fastap -SUFFIX=trpb -MINMATCH=0.1
-DOOPUL=0 -DOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=10sum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09890220.ecgn1.1.156.0runat.13062003.144854.16586
-NCPU=6 -ICPU=3 -NO_MMAR -LAAGREUERY -NEG_SCORES=0 -WAIT -DSPBLCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2.6/ptodata1/pubna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata1/pubna/PCCT_NEM_PUB.seq:*
3: /cgn2.6/ptodata1/pubna/US06_NEM_PUB.seq:*
4: /cgn2.6/ptodata1/pubna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata1/pubna/US07_NEM_PUB.seq:*
6: /cgn2.6/ptodata1/pubna/PCCT_PUBCOMB.seq:*
7: /cgn2.6/ptodata1/pubna/US08_NEM_PUB.seq:*
8: /cgn2.6/ptodata1/pubna/US08_PUBCOMB.seq:*
9: /cgn2.6/ptodata1/pubna/US09_NEM_PUB.seq:*
10: /cgn2.6/ptodata1/pubna/US09_PUBCOMB.seq:*
11: /cgn2.6/ptodata1/pubna/US10_NEM_PUB.seq:*
12: /cgn2.6/ptodata1/pubna/US10_PUBCOMB.seq:*
13: /cgn2.6/ptodata1/pubna/US60_NEM_PUB.seq:*
14: /cgn2.6/ptodata1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	423	17.8	2556	US-09-938-842A-2418
2	266	11.2	888	US-09-938-842A-453
3	225	9.3	3545	US-09-764-864-144
4	221	9.3	2217	US-09-874-162A-6

5	221	9.3	2328	US-09-874-162A-9	Sequence 9, Appl1
6	221	9.3	4409	US-09-874-162A-7	Sequence 7, Appl1
7	221	9.3	4441	US-09-874-162A-4	Sequence 4, Appl1
8	221	9.3	4441	US-09-969-347-169	Sequence 169, App
9	157.5	6.6	1257	US-09-822-830A-428	Sequence 428, App
10	153	6.4	2481	US-09-764-864-124	Sequence 124, App
11	129	5.4	4985	US-10-094-240-10	Sequence 35, Appl1
12	128	5.4	2001	US-09-801-368-35	Sequence 111, App
13	127.5	5.4	2886	US-09-801-368-131	Sequence 1, Appl1
14	124	5.2	3931	US-10-006-780-1	Sequence 472, App
15	123	5.2	264	US-09-969-373-472	Sequence 42, Appl1
16	122.5	5.2	3996	US-10-087-464-42	Sequence 568, App
17	115	4.8	526	US-09-764-864-568	Sequence 100, App
18	113.5	4.8	32768	US-09-070-927A-100	Sequence 11, Appl1
19	109	4.6	2237	US-09-994-485-7	Sequence 1872, App
20	109	4.6	1299	US-09-938-842A-1872	Sequence 143, App
21	103	4.3	3150	US-10-228-531-13	Sequence 129, App
22	103	4.3	6868	US-09-981-876-129	Sequence 129, App
23	103	4.3	2713	US-09-148-545-129	Sequence 123, App
24	101.5	4.3	2713	US-09-801-368-123	Sequence 582, App
25	101.5	4.2	1605	US-09-801-368-123	Sequence 81, Appl1
26	101	4.2	1145	US-10-114-170-212	Sequence 593, App
27	99.5	4.2	1145	US-09-764-864-582	Sequence 255, App
28	98.5	4.1	4775	US-10-146-477-81	Sequence 9834, App
29	98.5	4.1	7425	US-10-083-357-593	Sequence 163, App
30	97	4.1	1320	US-10-083-357-593	Sequence 3152, App
31	97	4.1	1680	US-09-801-368-255	Sequence 593, App
32	96.5	4.1	5085	US-10-198-846-9854	Sequence 255, App
33	96.5	4.1	6723	US-09-070-927A-163	Sequence 9834, App
34	96	4.0	333	US-09-938-842A-3152	Sequence 163, App
35	96	4.0	1956	US-09-351-794A-1	Sequence 3152, App
36	96	4.0	3282	US-09-801-368-391	Sequence 391, App
37	95	4.0	2098	US-10-286-264-91	Sequence 91, Appl1
38	95	4.0	2277	US-09-801-368-223	Sequence 223, App
39	95	4.0	9279	US-09-801-368-171	Sequence 171, App
40	94.5	4.0	496	US-10-155-533-5	Sequence 5, Appl1
41	94.5	4.0	951	US-10-155-533-10	Sequence 10, Appl1
42	94.5	4.0	978	US-10-155-533-11	Sequence 11, Appl1
43	94.5	4.0	978	US-10-155-533-12	Sequence 12, Appl1
44	94.5	4.0	1672	US-10-153-666-75	Sequence 75, Appl1
45	94.5	4.0	1678	US-10-153-666-75	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-938-842A-2418
Sequence 2418, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2418
LENGTH: 2556
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2418
Alignment Scores:

Pred. No.: 7.14e-41 Length: 2556
 Score: 423.00 Matches: 101
 Percent Similarity: 40.39% Conservative: 2
 Best Local Similarity: 39.61% Mismatches: 0
 Query Match: 17.79% Indels: 152
 DB: 9 Gaps: 2

US-09-890-220-2 (1-445) x US-09-938-842A-2418 (1-2556)

OY 43 Proserpheleuproactylaseantylrlyslleclvalalysargylargyls--- 61
 DB 2101 CCATCGTTTCTGCCAGATCTTGAACCTACAAATGGGCAAGGCAAGAAAGATAT 2160
 OY 61 GCGTTTCTTCTGAGTCTGACAGTCCACAGTATGATGCTTACTCTTAATATG 61
 DB 2161 GCGTTTCTTCTGAGTCTGACAGTCCACAGTATGATGCTTACTCTTAATATG 2220
 OY 62 -----Serargser 64
 DB 2221 AAGCTGATGAACATATTTATCTTGTGAGATATGACATATGAATGATGCTACAGATCT 2280
 OY 65 ThrGlymetvalvalpheasnyrlylaspCyasnasntThrLeuGlnLysThrGluval 84
 DB 2281 ACTGGATGTAGTTCACATTAAGATTTGTAATTAATACATTTACAAAGAACTGAA--- 2337
 OY 85 ArgGluaspCyserCyserProphecysSerMetLeuCyGlySerPheLysGlyLeuGln 104
 DB 2337 ----- 2337
 OY 105 PheHisLeuanserSerHisAspLeupheGluPheGluPheLysLeupheGluGlyTyr 124
 DB 2337 ----- 2337
 OY 125 GlnThrValasnsvalserValLysLeuanserPheLlePheGluGluGlySerasp 144
 DB 2337 ----- 2337
 OY 145 AspAspLysPheGluPheProPheSerLeuCySerLysProArgLysArgGlnArgGly 164
 DB 2337 ----- 2337
 OY 165 GlyArgasnsnThrArgArgLeuLysValCyPheLeuProLeuAspSerProSerLeu 184
 DB 2337 ----- 2337
 OY 185 ThrAsnGlyThrGluAsnGlyLleThrLeuLeuansaspGlyAsnArgGlyLeuGlyTyr 204
 DB 2338 -----GGAACCGTGTTTATGAGATAT 2358
 OY 205 ProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnLleProProAlaLle 224
 DB 2359 CCGAGGACACAGAGCTGCTGACAAATTTGAGATGACATGACATTCGCCAGCCATA 2418
 OY 225 AlaHisSerLeuaspAlaGlyAlaLysValLleLeuThrSerGluAlaValaPro 244
 DB 2419 GCCCACTCTTCTGAGACCGTGTGCTAAAGTTAATTAACAACGAGCTGTGCTCCT 2478
 OY 245 AlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
 DB 2479 GCTACTAAGACAGAAAGTTATCTGCTGAGCATGAGAGGCTAGA 2523

RESULT 2

; Sequence 453, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Yun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO: 453
 ; LENGTH: 888
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana

US-09-938-842A-453

Alignment Scores:

Pred. No.: 1.93e-22 Length: 888
 Score: 266.00 Matches: 70
 Percent Similarity: 55.73% Conservative: 30
 Best Local Similarity: 53.44% Mismatches: 28
 Query Match: 11.19% Indels: 28
 DB: 9 Gaps: 4

US-09-890-220-2 (1-445) x US-09-938-842A-453 (1-888)

OY 278 MetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAla 297
 DB 1 ATGGCGCTTGAGCAGCATATCTGTGATCGGAGATAGCAGATGATGATGATGATGAT 60
 OY 298 AlaAspPheGluAspArgGln-----MetLeuAspAspPheValAspAlaAspGluGln 304
 DB 61 GCAGATTTTGAAGATCGCAGCATATTCATGATTTCTTCTGCTTATTAGTAGGCA 120
 OY 305 -----MetLeuAspAspPheValAspAlaAspGluGln 318
 DB 121 CAGAAATGCTATGATATATCTGATGATCTTGTGATGATGATGATGATGATGATGAT 180
 OY 319 PheMetHisLeuThrPheSerPheValArgLysGlnArgValLleAlaAspGlyHisLle 338
 DB 181 TTCATGCAATCTTGTGAACCTGTTGTAAAGAAACAAAGA-----AACAAACATATC 231
 OY 339 SerTTPAlaCyGluAlaPheSerArgPheThrGluGluLeuHisArgTyrSerSer 358
 DB 232 TTATTACCAATCTTGTCTTCATAGCAGATCCTCTCGACACCTTACAC-----CTCTCC 285
 OY 359 LeuphetPcystrPargLeupheLleLysLeuThrPasnHisGlyLeuValAspSer 378
 DB 286 TTA-----ACCTTACCTCTTCTGCTTAAACCAACATGCTTTGTTAGAA 333
 OY 379 AlaThrLleAsnAsnCyAsnThrLleLeuGlu 389
 DB 334 GCCTTAGCTAAGCTTGTGATACACAGGTTTGA 366

RESULT 3

; Sequence 144, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; PRIOR FILING DATE: 2001-01-17
 ; PRIOR APPLICATION data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SEQ ID NO: 144
 ; LENGTH: 3545
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-09-764-864-144

Alignment Scores:

Pred. No.: 1.99e-16 Length: 3545
 Score: 225.00 Matches: 84
 Percent Similarity: 37.09% Conservative: 64
 Best Local Similarity: 21.05% Mismatches: 151
 Query Match: 9.46% Indels: 100
 DB: 10 Gaps: 12

US-09-890-220-2 (1-445) x US-09-764-864-144 (1-3545)

QY 58 ArgLysArgLys-----SerArgSerThrGlyMetValValPheAsnTyr 72
 DB 503 AGAAAGAAAAGAGTACTCCAAATGAAAACCGACAAAATTAAGAAATTTTATACAGTTT 562
 QY 73 LysAspCysAsnThrLeuGlnLysThrGluValArgLysAspCysSerCysProPhe 92
 DB 563 CTCTATTAACAACATATACAGCAACAACCTGAACCAAGACATGACCTGCAATGCCCTTGG 622
 QY 93 CysSerMetLeuGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
 DB 623 TGTACTCTGAACCTGCCCAAACTTTATAGTTTACTCAACGATCTTAACCTGCGCATAGC 682
 QY 113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
 DB 683 AGATTATCTTCACTATGTTTATCATCCAAAAGCTGAGATGATGTTTCTATC--- 739
 QY 133 LeuAsnSerPheIlePheGluGluGlySerAspAspAspLysPheGluProPheSer 152
 DB 740 -----AATGAGGTTATGATGCG-----TTC 760
 QY 153 LeuCysSerLysProArgLys---ArgArgGlnArgGly-----GlyArgAsnAsn 168
 DB 761 TATGACAGAAATCTTCAGATATTCACGCCAACCTGATGCTTTAGTCCGACAGCA 820
 QY 169 ---ThrArgArgLeuGlyValLysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
 DB 821 CCAAGTTAAGAAACACCTATCACACATATTTCTGTGCGACGCGCAAAAACGACAAAAGCA 880
 QY 188 ThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
 DB 881 AGCATGTCGATATTTCTGTAATCTGAAGATGCG----- 913
 QY 208 ThrGluLeuAlaGlyLysPheGluMetThrSerAsnIleProAlaIleAlaHisSer 227
 DB 913 ----- 913
 QY 228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValProAlaThrLys 247
 DB 913 ----- 913
 QY 248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArg 267
 DB 914 -----GAAGTAGACGCAAAAGACATATAGTAGTGCCAC-----AATCGT 955
 QY 268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGlnGluValMetSerAspArg 287
 DB 956 CTGTATTTCCATAGAGATCCGCTTACCTCCGCTCCGACCAAGAAATG-----GAAGTA 1009
 QY 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
 DB 1010 GAATAGTGACAGATGAAAAGATCTGAATGCTTAAGACAAAACCATTTACACAAATTGAA 1069
 QY 308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrPAsnSerPheVal 327
 DB 1070 GAATTTCGTATGTATGAGAGAGACAAAGATGATGAACCTGTGGAATCTTCATAGTC 1129
 QY 328 ArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSerArg 347
 DB 1130 ATGAAGCATGGTTATTTCTGACATCAATCAATGATCATGCTGTATGCTTTGTAGAA 1189
 QY 348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheThrPyrCysThrPArgLeuPheLeu 367
 DB 1190 AATTATGGACAGAAATTAATTAG---AAGAATTTATGTGCAAACTTCATGCTTCATCTA 1246

QY 368 IleLysLeuThrPAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
 DB 1247 GTCAGCATGATGATGACTTATATCTTATAGCATATATGATCATAGTAAAGCTGTACAGAA 1306
 QY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAsp 407
 DB 1307 CTCCTGTAAGATGACG-----CAAAAATTGAA 1333
 QY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
 DB 1334 AAGGGGAATCTGCTTCCCTCGCAACAGAAATTACTGAAGAAACAAATGGGACAGCA 1393
 QY 425 -----AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer 439
 DB 1394 AATGATTTAGTGAATTAATTAATCTCAAAAGACAAAGCTTTGGAACAGATAGTCTCA 1450

RESULT 4

US-09-874-162A-6
 ; Sequence 6, Application US/09874162A
 ; Patent No. US20020155452A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Koonitz, Jason
 ; APPLICANT: Sklar, Jeffrey
 ; TITLE OF INVENTION: FUSION OF JAZF1 AND JAZF1 GENES IN
 ; TITLE OF INVENTION: ENDOMETRIAL STROMAL TUMORS
 ; FILE REFERENCE: 05311-024001
 ; CURRENT APPLICATION NUMBER: US/09/874,162A
 ; CURRENT FILING DATE: 2001-06-04
 ; PRIOR APPLICATION NUMBER: US 60/209,093
 ; PRIOR FILING DATE: 2000-06-02
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 2217
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-874-162A-6

Alignment Scores:

Pred. No.: 2.81e-16 Length: 2217
 Score: 221.00 Matches: 84
 Percent Similarity: 37.09% Conservative: 64
 Best Local Similarity: 21.05% Mismatches: 151
 Query Match: 9.29% Indels: 100
 DB: 9 Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-6 (1-2217)

QY 58 ArgLysArgLys-----SerArgSerThrGlyMetValValPheAsnTyr 72
 DB 1237 AGAAAGAAAAGAGTACTCCAAATGAAAACCGACAAAATTAAGAAATTTTATACAGTTT 1296
 QY 73 LysAspCysAsnThrLeuGlnLysThrGluValArgLysAspCysSerCysProPhe 92
 DB 1297 CTCTATTAACAACATATACAGCAACAACCTGAACCAAGACATGACCTGATGCTTGG 1356
 QY 93 CysSerMetLeuGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
 DB 1357 TGTACTCTGAACCTGCCCAAACTTTATAGTTTACTCAAGATCTTAACCTGCGCATAGC 1416
 QY 113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
 DB 1417 AGATTATCTTCACTATGTTTATCATCCAAAAGCTGAGATGATGTTTCTATC--- 1473
 QY 133 LeuAsnSerPheIlePheGluGluGlySerAspAspAspLysPheGluProPheSer 152
 DB 1474 -----AATGACTGTTATGATGCG-----TTC 1494
 QY 153 LeuCysSerLysProArgLys---ArgArgGlnArgGly-----GlyArgAsnAsn 168
 DB 1495 TATGACAGAAATCTTCAGATATTCACGCCAACCTGATGCTTTAGTCCGACAGCA 1554
 QY 169 ---ThrArgArgLeuLysValLysPheLeuProLeuAspSerProSerLeuThrAsnGly 187

```

DB 1555 CCAGTTAGAGAACACCTATCAGACATATCTTGTGTGCGAGCGCAAAAGCAAAAGCA 1614
      |||||
QY 158 ThrGluasnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
      |||||
DB 1615 AGCATGCTGAAATTTCTTGATCTGAAGATGG----- 1647
QY 208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227
      |||||
DB 1647 ----- 1647
QY 228 SerLeuAspAlaGlyAlaValIleLeuThrSerGluAlaValProAlaThrIlys 247
      |||||
DB 1647 ----- 1647
QY 248 ThrArgIleuSerAlaGluArgSerGluAlaArgSerHisIleuLeuGlnIlysaArg 267
      |||||
DB 1648 -----GAAGTAGAACAGCAAGAACATATAGTAGGCCAC-----AATCGT 1689
QY 268 GlnPheTyrHisSerHisArgValGlnPrometAlaLeuGlnGlnValMetSerAspArg 287
      |||||
DB 1690 CTGATATTCATAGTACCTGCTGCTCTCCGCCACAGAAATG-----GAAGTA 1743
QY 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
      |||||
DB 1744 GATAGTGAAATGAAAGATCCGATGCGTAAAGAAACCATTTACACAAATTGAA 1803
QY 308 AspPheValAspValAsnIlyAspGlyGlnPheMetHisIleuThrAsnSerPheVal 327
      |||||
DB 1804 GAGTTTCTGATGTATTAAGAGAGAGAAAGATGATGAACCTCGAATCTCCATGTC 1863
QY 328 ArgIlyGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGlnAlaPheSerArg 347
      |||||
DB 1864 ATGAAGCATGGGTTATTCCTGACAAATCAATCAATCAATCCCTGTTGCTTTGTAGAA 1923
QY 348 PheTyrGluIlyGlnLeuHisArgTyrSerSerLeuPheThrPcysThrArgLeuPheLeu 367
      |||||
DB 1924 AATTATGACAGAAATTAATTAAG--AGAAATTTATGCGAAATCTTCGCTCATCTCA 1980
QY 368 IleIlyIleuThrAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
      |||||
DB 1981 GTCAAGCATGACATCTTAATCTTAATGACATATGATTAAGATTAAGCTTTTACCAAG 2040
QY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAsp 407
      |||||
DB 2041 CTCGGTGAAATGCG-----CAAAATTTAGAA 2067
QY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
      |||||
DB 2068 AAGGGGGAATCTGCTCCCTGCAGAAAGAAATTAAGAAATGAAACAAATGGACAGCA 2127
QY 425 -----AspIleAsnAsnIlyAsnAsnValAspAsnIlyAspAsnAsnSer 439
      |||||
DB 2128 AATGATTTAGTGAAATTAACCTCAAAAGAGAAAGCTTTGAAACAGATGTGTCTCA 2184
      |||||

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-874-162A-9

Alignment Scores:
Score: 221.00
Percent Similarity: 37.09%
Best Local Similarity: 21.05%
Query Match: 9
Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-9 (1-2328)

QY 58 ArgIlyAspArgIly-----SerArgSerThrGlyMetValIlePheAsnTyr 72
      |||||
DB 1348 AGAAAGAAAGAAAGATCTCAAAATGAAACCGCAAAATTAAGAAATATTTATATCAAGTTT 1407
QY 73 LysAspCysAsnAsnThrLeuGlnIlySerThrGluValAlaArgGluAspCysSerCysProPhe 92
      |||||
DB 1408 CTCTATTAACAAACAAATACAGGACACAAACTGAACCAAGATGACATGCTGCTGCTGG 1467
QY 93 CysSerMetLeuCysGlySerPheIlyGlyGlnPheHisIleuAsnSerSerHisAsp 112
      |||||
DB 1468 TGTACTCTGACACTGCGCAACTTATATGTTTACTCAAGCATCTTAACTGCTGCATAC 1527
QY 113 LeuPheGluPheGluPheIlyLeuPheGluIlyGlyGlnThrValAsnValSerValIys 132
      |||||
DB 1528 AGATTATCTTCAACTATGTTTATCATCCAAAGCTGCTAGATATGTTTCTATC----- 1584
QY 133 LeuAsnSerPheIlePheGluGlnGlyIlySerAspAspIlySerPheGluProPheSer 152
      |||||
DB 1585 -----AATGAGCTTATGATGCG-----TCC 1605
QY 153 LeuCysSerIlyProArgIly-----ArgArgGlnArgIly-----GlyArgAsnAsn 168
      |||||
DB 1606 TATGCAAGAAATCTCAGAGATATTCATGCGCAACCTGATGCTTTTACTGCGCAACGCA 1665
QY 169 ---ThrArgArgLeuIlyValCysPheIleuProLeuAspSerProSerLeuThrAsnGly 187
      |||||
DB 1666 CCAGTTAAGAACACCTATCATCACAAATTTCTGTGCGAGCGCAAAACGCAAAAGCA 1725
QY 188 ThrGluAsnGlyIleThrLeuLeuAsnAspIlyAsnArgGlyLeuGlyTyrProGluAla 207
      |||||
DB 1726 AGCATGCTGCAATTTCTTGAATCTGAAGATGG----- 1758
QY 208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227
      |||||
DB 1758 ----- 1758
QY 228 SerLeuAspAlaGlyAlaValIleLeuThrSerGluAlaValProAlaThrIlys 247
      |||||
DB 1758 ----- 1758
QY 248 ThrArgIlyLeuSerAlaGluArgSerGluAlaArgSerHisIleuLeuGlnIlysaArg 267
      |||||
DB 1759 -----GAAGTAGAACAGCAAGAACATATAGTAGGCCAC-----AATCGT 1800
QY 268 GlnPheTyrHisSerHisArgValGlnPrometAlaLeuGlnGlnValMetSerAspArg 287
      |||||
DB 1801 CTGATTTTCCATAGATGACCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1854
QY 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
      |||||
DB 1855 GATAGTAGAACAGAAAGATCTGAAATGCGTAAAGAAACAAATTTACAAATTTGAA 1914
QY 308 AspPheValAspValAsnIlyAspGlyGlnPheMetHisIleuThrAsnSerPheVal 327
      |||||
DB 1915 GAGTTTCTGATGTATTAAGAGAGAAATGATGATGAAATCTGCAATGCTCATGTC 1974
QY 328 ArgIlyGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSerArg 347
      |||||
DB 1975 ATGAAGCATGGGTTATTTATGTCACATCAATGATATGCTGTGTTGTAGAA 2034
      |||||

```

```

RESULT 5
US-09-874-162A-9
; Sequence 9, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: Koontz, Jason
; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAZF1 AND JAZ1 GENES IN
; TITLE OF INVENTION: ENDOMETRIAL STROMAL TUMORS
; FILE REFERENCE: 05311-024001
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,093
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2328

```



```

OY 348 PheTYrGluYsGluLeuHnIsArGTySerSerLeuPheTrpCysTrpAlaPheLeu 367
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2033 AATTATGACAGAAATAATTAG--AAGATTATATGTCGAATCTTCATCTCA 2091
OY 368 IleYsLeuTrpAsnHsIsgLYeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2092 GTCAGATGCATGACTTAACTTATTTATGCAATAGTCAATAGATTAACCTGTTCACAG 2151
OY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrasnAsnAsnSerValAsp 407
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2152 CTCCTGGAATGCAG-----CAAAATATAGAA 2178
OY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProsn----- 424
DB 2179 AAGGGGAATCTGCTTCCCTCCCAACGAAAGAAATAGTGAAGAACAAATAGGACACGA 2238
OY 425 -----AspIleAsnAsnLYsAsnAsnValAspAsnLYsAsnAsnSer 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2239 AATGATTTAGTGAATTAATCAAAAGAAAGCTTTGAAACAGATAGTGTCTCA 2295

RESULT 6
US-09-874-162A-7
; Sequence 7, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: Koontz, Jason
; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAF1 AND JAZ1 GENES IN
; FILE REFERENCE: 05311-024001
; CURRENT APPLICATION NUMBER: US/09/874,162A
; CURRENT FILING DATE: 2001-06-04
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(2379)
US-09-874-162A-7

Alignment Scores:
Pred. No.: 8,83e-16 Length: 4409
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
DB: 9 Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-7 (1-4409)
OY 58 ArgLYsArgLYs-----SerArgSerThrGLYmetValValPheAsnTYr 72
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1399 AGAAAGAAAGAAAGATCTCAAAATGAACCGCAAAATAATTAAGAAATATTTATCAAGTTT 1458
OY 73 LysAspCysAsnAsnThrLeuGluInLYsThrGluValArgLYsAspCysSerCysProPhe 92
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1459 CTTCTAATACACATATCAAGGACACAACTGAAGCAAGAGATGACCTGATTCGCTTGG 1518
OY 93 CysSerMetLeuCysGLYSerPheLYsGLYLeuGlnPheHisLeuAsnSerSerHisAsp 112
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1519 TGTACTCTGAACTGCGCAAACTTATAGTTTACTCAACACACTTAACTGTGCATAGC 1578
OY 113 LeuPheGluPheGluPheLYsLeuPheGluInLYsThrGlnThrValAsnValSerValLYs 132
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1579 AGATTATCTCAACTATATTTATATCATCCAAAGGCTGACGATAGATGTTCTATATC--- 1635
OY 133 LeuAsnSerPheIlePheGluInLYsLeuAspAspLYsPheGluPhePropheSer 152

```

```

DB 1636 -----AATGAGTGTATAGAGGC-----TTC 1656
OY 153 LeuCYsSerLYsProArgLYs---ArgArgGlnArgLYs-----GlyArgAsnAsn 168
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1657 TATGACGAAATCTCAGGATATTCATCCCAACCGAGATTTGCTTTAGTGTACACGGA 1716
OY 169 ---ThrArgArgLeuLYsValCYsPheLeuProLeuAspSerProSerLeuThrAsnGLY 187
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1717 CCAGTTAAGAGAACCTATTCACACATATTTCTGTGTGCAGGCGCAAAACGAAACAAAGCA 1776
OY 188 ThrGluAsnGLYIleThrLeuAsnAspGLYAsnArgLYsLeuGLYTrpProGluAla 207
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1777 AGCAGTCTGAATTTCTTGAATCTGAAGATGGG----- 1809
OY 208 ThrGluLeuAlaGLYInPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227
DB 1809 ----- 1809
OY 228 SerLeuAspAlaGLYAlaLYsValIleLeuThrSerGluAlaValProAlaThrLYs 247
DB 1809 ----- 1809
OY 248 ThrArgLYsLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGluInLYsArg 267
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1810 -----GAAGTGAACAGCAAGAAACATATAGTAGGCGCAC-----AATCGT 1851
OY 268 GlnPheTYrHisSerHisArgValGlnProMetAlaLeuGluInValMetSerAspArg 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1852 CTGTATTTCCAAATGATGATCCCTGCTTACCTCCGCTCCACAGAAATG-----GAAGTA 1905
OY 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1906 GATAGGAAAGATGAAGAAAGATCTGATGCTAAGGAAACCACTTACACAAATGGA 1965
OY 308 AspPheValAspValAsnLYsAspGluLYsGlnPheMetHisLeuTrpAsnSerPheVal 327
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1966 GAGTTTTCGAGTTAATGATGAGAGAAAGAAAGATGATGAACCTGGAATCTCCATGTC 2025
OY 328 ArgLYsGlnArgValIleAlaAspGlnHisIleSerTrpAlaCysGluAlaPheSerArg 347
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2026 ATGAAAGCATGGTTATGTTCTGTCACATCAAAATGATCATGCTGTGTTGTAGAA 2085
OY 348 PheTYrGluLYsGluLeuHnIsArGTySerSerLeuPheTrpCysTrpAlaPheLeu 367
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2086 AATTATGACAGAAATAATTAG--AAGATTATATGTCGAATCTTCATCTCA 2142
OY 368 IleYsLeuTrpAsnHsIsgLYeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2143 GTCAGATGCATGACTTAACTTATTTATGCAATAGTCAATAGATTAACCTGTTCACAG 2202
OY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrasnAsnAsnSerValAsp 407
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2203 CTCCTGGAATGCAG-----CAAAATATAGAA 2229
OY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProsn----- 424
DB 2230 AAGGGGAATCTGCTTCCCTCCCAACGAAAGAAATAGTGAAGAACAAATAGGACACGA 2289
OY 425 -----AspIleAsnAsnLYsAsnAsnValAspAsnLYsAsnAsnSer 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2290 AATGATTTAGTGAATTAATCAAAAGAAAGCTTTGAAACAGATAGTGTCTCA 2346

RESULT 7
US-09-874-162A-4
; Sequence 4, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: Koontz, Jason
; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAF1 AND JAZ1 GENES IN
; FILE REFERENCE: 05311-024001

```

```

CURRENT APPLICATION NUMBER: US/09/874,162A
CURRENT FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: US 60/209,093
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 4441
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (195)...(2411)
US-09-874-162A-4

Alignment Scores:
Pred. No.: 8,94e-16 Length: 4441
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-4 (1-4441)
OY 58 ArglysaArglys-----SerArgSerThrglymetValaIphesantyr 72
Db 1431 AGAAAAGAAAAGGATCTCCAAATGAAACCGCAAAATTAAGAAATTTATTCAGTTT 1490
OY 73 LysaspCysasnAnthrleuglnlyThrgluValaArggluAspCysSerCysProphe 92
Db 1491 CTCATATACAAACATACAGCAAGCAACAACTGAAGCAAGATGACCTTCATGCCCTTG 1550
OY 93 CysSerMetleuCySglYserPheylsGlyleuGlnPheHisLeuAsnSerSerHisasp 112
Db 1551 TGTACTCTGAACCTCCGCAAACTTTATCTTTCACAGCACTTAATCTGCCATAGC 1610
OY 113 leuphegluPhegluPheylsleuphegluGluTyrGlnThValaAsnValaSerVallys 132
Db 1611 AGATTATCTTCAACTTGTGTTTATCTCCAAAGGCTGAGATGATGTTCTATTC- 1667
OY 133 LeuAsnSerPhelelePhegluGluGlySerAspAspAspPhegluProPheSer 152
Db 1668 -----AATGAGTGTATGATGCG-----TCC 1688
OY 153 leuCySerlySProArglys---ArgArgGlnArggly-----GlyArgAsnAsn 168
Db 1689 TATGCAAGAAATCTCCAGATATTCATCCCAACCTGGATTTGCTTTAGTCGCAAGCGA 1748
OY 169 ---ThrArgArglyleuValaCysPheleuProleuAspSerProSerleuThrAsnGly 187
Db 1749 CCAAGTAAAGAAACACCTATTCACACATATTTCTGTGTCAGGCCCAAAAGCAAAAGCA 1808
OY 188 ThrGluAsnGlyleThrleuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
Db 1809 AGCATGCTCATATTTCTGAATCTGAATCGAATGGG----- 1841
OY 208 ThrGluLeuAlaGlyGlnPhegluMetThrSerAsnIleProProAlaIleAlaHisSer 227
Db 1841 ----- 1841
OY 228 SerleuAspAlaGlyAlaValaValIleleuThrSerGluAlaValaIleProAlaThrlys 247
Db 1841 ----- 1841
OY 248 ThrArgGlyleuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArg 267
Db 1842 -----GAAGTAAAGCAAGCAAAACATATCTACTATGCGCAC-----AATCGT 1883
OY 268 GlnPheYthrHisSerHisArgValaGlnProMetAlaLeuGluGlnValMetSerAspArg 287
Db 1884 CTGTATATTCATAGTATGATACCTGCTTACCTCCGTCACAAAGAAATG-----GAAGTA 1937

```

```

OY 288 AspSerGluAspGluValaIAspAspAspValaIAspPhegluAspArgGlnMetLeuAsp 307
Db 1938 GATAGTAAAGTAAAGGATCTCCAAATGAAACCGCAAAATTAAGAAATTTATTCAGTTT 1490
OY 308 AspPheValaSpValaAsnLysAspGluGlnPheMetHisLeuThrAsnSerPheVal 327
Db 1998 GAGTTTCTGATGTTAAAGAGAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2057
OY 328 ArglyGlnArgValaIleAlaAspGlyHisIleSerThrAlaCysGluAlaIleSerArg 347
Db 2058 ATGAAGCATAGGTTTATTTATTCGACAAATCAATGAAATGATGCTGTATGCTGTATGAA 2117
OY 348 PheYrglyGlyGluLeuHisArgTyrSerSerleuPheThrPcysThrArgleuPheleu 367
Db 2118 AATTATGACAGCAAGAAATATTAAG-----AAGAAATTTATGTCGAAACCTTCATGCTATCTA 2174
OY 368 IleYsleuThrPasnHisGlyLeuValaIAspSerAlaThrIleAsnAsnCysAsnThrIle 387
Db 2175 GTCACACATGACATGACTTATATCTTATAGCATATGATGATGATGATGATGATGATGATG 2234
OY 388 leuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnSerValaIAsp 407
Db 2235 CTCGCTGAATGACAG-----CAAAATTTAGAA 2261
OY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValaIAspHisProAsn----- 424
Db 2262 AAGGGGGAATGCTGCTCCCGCAAGCAAGAAATTAAGTGAACAAATGGACAGCA 2321
OY 425 -----AspIleAsnAsnLysAsnAsnValaIAspAsnLysAspAsnSer 439
Db 2322 AATGATTTATGTAATTAATCTCAAAAGAGAAACCTTTGGAACAGATAGTGTCTCA 2378

RESULT 8
US-09-969-347-169
? Sequence 169, Application US/09969347
? Patent No. US20020115085A1
? GENERAL INFORMATION:
? APPLICANT: Eber, Reinhard
? TITLE OF INVENTION: Sets
? FILE REFERENCE: 689290-69
? CURRENT APPLICATION NUMBER: US/09/969,347
? PRIOR FILING DATE: 2000-10-02
? PRIOR APPLICATION NUMBER: US/60/237,598
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: US/60/237,604
? NUMBER OF SEQ ID NOS: 318
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 169
? LENGTH: 4441
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-969-347-169

Alignment Scores:
Pred. No.: 8,94e-16 Length: 4441
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
Gaps: 12

US-09-890-220-2 (1-445) x US-09-969-347-169 (1-4441)
OY 58 ArglysaArglys-----SerArgSerThrglymetValaIphesantyr 72
Db 1431 AGAAAAGAAAAGGATCTCCAAATGAAACCGCAAAATTAAGAAATTTATTCAGTTT 1490
OY 73 LysaspCysasnAnthrleuglnlyThrgluValaArggluAspCysSerCysProphe 92
Db 1491 CTCATATACAAACATACAGCAAGCAACAACTGAAGCAAGATGACCTTCATGCCCTTG 1550

```

```

QY 93 CysSerMetLeuCysGlySerPheGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
    ||| ||| ||| ||| |||
Db 1551 TGTACTCTGAACTGGCGCAAACTTTATTACTTACCAACATCTTAAATCTGGCCATTAC 1610
QY 113 LeuPheGluPheGluPheGlyLeuPheGluGlyTyrGlnThrValAsnValSerValLys 132
    ||| ||| ||| ||| |||
Db 1611 AGATTATTCCTCAACATGTTTATTCATCCAAAGAGTCAGATACATGTTCTTATC--- 1667
QY 133 LeuAsnSerPheLeuPheGluGluGlySerAspAspAspLysPheGluProPheSer 152
    ||| ||| ||| ||| |||
Db 1668 -----AATGAGTCTTATGATGGC-----TCC 1688
QY 153 LeuCysSerLysProArgLys---ArgArgGlnArgGly-----GlyArgAsnAsn 168
    ||| ||| ||| ||| |||
Db 1689 TATGCAGAAATCCCTCAGATATTCATGCCCAACCGATTTGCTTTAGTGCACAGCA 1748
QY 169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
    ||| ||| ||| ||| |||
Db 1749 CCAGTTAAGAGAACACCTATCACACATATTCCTGTCGACGCCAAACGAAACAAAGCA 1808
QY 188 ThrGluAsnGlyLeuThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
    ||| ||| ||| ||| |||
Db 1809 ACATGCTCGAATTTCTTGAATCTGAAGATGGG----- 1841
QY 208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227
    ||| ||| ||| ||| |||
Db 1841 ----- 1841
QY 228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValAlaProAlaThrLys 247
    ||| ||| ||| ||| |||
Db 1841 ----- 1841
QY 248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArg 267
    ||| ||| ||| ||| |||
Db 1842 -----GAAGTGAACAGCAAGAAACATATAGTAGGCCAC-----AATCCGT 1883
QY 268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGlnGlnValMetSerAspArg 287
    ||| ||| ||| ||| |||
Db 1884 CTTGATTTCCATAGTAGTACCTGCTTACTCTCCGCCACAGAAATG-----GAAGTA 1937
QY 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
    ||| ||| ||| ||| |||
Db 1938 GATAGAGAAATGAAAGAAAGCATCTGATAGGAAAGAAACCAATTCACAAATGTGA 1997
QY 308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPheVal 327
    ||| ||| ||| ||| |||
Db 1998 GAGTTTCTGATGTTATGAGAGAGAGAAAGATGATGAATCTGGAATCTCCATGTC 2057
QY 328 ArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSerArg 347
    ||| ||| ||| ||| |||
Db 2058 ATGAGAGCATGGGTTTATGCTGACATCAATGAATCAATGCCGTGTTTGTAGAA 2117
QY 348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheThrPcysThrArgLeuPheLeu 367
    ||| ||| ||| ||| |||
Db 2118 AATTATGACAGAAATAATTAG---AAGAATTTATGCGAAACTTCATGCTTCATCTTA 2174
QY 368 IleLysLeuThrAsnHisGlyLeuValAspSerAlaThrIleAsnAsnLysAsnThrIle 387
    ||| ||| ||| ||| |||
Db 2175 GTCAGCATCATGATCTTATATCTTATTACATATATGTCATATGATTAACCTGTTACCAAG 2234
QY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnSerValAsp 407
    ||| ||| ||| ||| |||
Db 2235 CTCCTCGAANAATGAG-----CAAAATATAGAA 2261
QY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
    ||| ||| ||| ||| |||
Db 2262 AAGGGGGAATCTGCTTCCCTCCGCAAGAGAAATTAATCAAGAACAAATGGGACAGCA 2321
QY 425 -----AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSer 439
    ||| ||| ||| ||| |||
Db 2322 AATGATTTATGATTAATTAATCAAGAAAGAAAGCTTTGGAACAGATAGTGTCTCA 2378

```

RESULT 9

```

US-09-822-830A-428
; Sequence 428, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 428
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 860
; OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-428

Alignment Scores:
Pred. No.: 6,36e-09 Length: 1257
Score: 157.50 Matches: 54
Percent Similarity: 45.658 Conservative: 51
Best Local Similarity: 23.488 Mismatches: 103
Query Match: 6.628 Indels: 23
DB: 10 Gaps: 5

US-09-890-220-2 (1-445) x US-09-822-830A-428 (1-1257)
QY 217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyValIle 236
    ||| ||| ||| ||| |||
Db 146 ACGAGCAGCATTAAGAGAACACCTATCATCATCTTGTGTCAGGCCA-AAACGAGCA 204
QY 237 LeuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluSer 256
    ||| ||| ||| ||| |||
Db 205 AAAGCAAGCATGCTGTAATTTCTGAATCTGAAGATGGGAGATAGAACAGCAAGACAA 264
QY 257 GluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
    ||| ||| ||| ||| |||
Db 265 TATAGTAGGGCCAC-----AATCGCTGATTTTCATATGATGATACCTCTTA 312
QY 277 PrometAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAsp 296
    ||| ||| ||| ||| |||
Db 313 CCTCTCGCCACAAAGAAATG-----GAAGTAGATAGGAAGATGAAAGATCCTGAA 366
QY 297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
    ||| ||| ||| ||| |||
Db 367 TGGCTAAGAGAAAACCATTAACAAATTTGAAGAGATTTTCTGATGTTAATGAGAGAG 426
QY 317 LysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
    ||| ||| ||| ||| |||
Db 427 AAAGAAATGATGAACCTCTGGAATCTCCATGTCATAGACATAGGGTTTATGCTGCAAT 486
QY 337 HisIleSerThrAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356
    ||| ||| ||| ||| |||
Db 487 CAATGAATCATGCTGATGCTGTTGTAGAAAATTAATGAGCAAGAAATAATTAATTAAG 543
QY 357 SerSerLeuPheThrPcysThrArgLeuPheLeuIleLysLeuThrAsnHisGlyLeuVal 376
    ||| ||| ||| ||| |||
Db 544 AAGAAATTTATGTCGAAACATTCATGCTTCACTAGTCAGATGATGATGATTAATCTTATT 603

```

OY 377 AsperalathrilleasnasncysasnThilleu----- 388
 DB 604 AGCATATGTCATTAAGTAAAGCTTTACCACTCCGTGAATATGACGAAATAATAGAA 663
 OY 389 -----GluasnCysArgasnSerSerAspThrThrThrasnasnasnSerVal 406
 DB 664 AAGGGGAAATCTGCTCCCTCCCTGCAAGAAATACTGAGAACMAAATGCGACACA 723
 OY 407 AspargProSerAspSerAsnThr-----AsnasnasnThleValasp 421
 DB 724 AATGATTTACTGAAATTAACCTCAAAAGAAAGCTTTGGAAACAGATAGTCTCAGCG 783
 OY 422 HisProsnAspIleasnAsnLysasnAsp 431
 DB 784 TTTCAAAAACAGACGCAAAAAAACAACTCT 813

RESULT 10

US-09-764-864-124
 ; Sequence 124, Application us/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; PRIORITY FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1782
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 124
 ; LENGTH: 2481
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-864-124

Alignment Scores:

Pred. No.: 7e-08 Length: 2481
 Score: 153.00 Matches: 53
 Percent Similarity: 44.35% Conservative: 49
 Best Local Similarity: 23.04% Mismatches: 103
 Query Match: 6.43% Indels: 26
 DB: 10 Gaps: 5

US-09-890-220-2 (1-445) x US-09-764-864-124 (1-2481)

OY 217 ThreSerasnIleProProlaIleAlaHisSerSerLeuAspAlaGlyAlaValValIle 236
 DB 200 ACGGACGAGTTAAGAGAACCTATCACACATATCTGTGTGCGAGGCGA-AAACGAAACA 258
 OY 237 LeuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
 DB 259 AAACGACAGCAGTCGTGAATTTCTTGAATCTGAAGGGGAGTGAACACGAAAGAACCA 318
 OY 257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
 DB 319 TATAGTAGTGGCCAC-----AATGCTGTGATTTCCATAGTGAATACCTGCTTA 366
 OY 277 PrometAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
 DB 367 CCTCTCGCTCCACAGAAATG-----GAAGTAGATAGTGAAGATGAAGAGATCCCTGAA 420
 OY 297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
 DB 421 TGGCTAAGAGAAAAACCATTAACAATAATTGAAGAGTTTTCGATGTTAATGAAGAGAG 480
 OY 317 LysGlnPheMetHisLeuThrPAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
 DB 481 AAAGAACTGTGAACCTGGAATCTCCATGTCAAGCAAGCATGGTATTATGCTGACAT 540
 OY 337 HisLeuSerThrAlaGluAlaPheSerArgPheTyrGluLysGlnLeuHisArgTyr 356
 DB 541 CAATGAAATCATGCTGTGATAGCTTTGTAGAAAAATTATAGCAGAAAAATATATAG--- 597

OY 357 SerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuThrPAsnHisGlyLeuVal 376
 DB 598 AAGAAATTTATGTGGAACCTGATGCTCATCTCATCTGATGATGATCACTTAATCTTAAT 657
 OY 377 AsperalathrilleasnasncysasnThrilleuGluasnCysArgasnSerSerVal 396
 DB 658 AGCATATGTCATTAAGTAAAGCTTTACCACTCCGTGAATATGACGAAATAATAGAA 705
 OY 397 ThrThrThrThrasnasnasnSerValAspArgProSerAspSerAsnThrAsnAsn 416
 DB 706 -----CAAAATTTGAAGAGGGAATCTGCTCCCTCGCAAC 744
 OY 417 AsnasnIleValAspHisProasn-----AspIleAsnasnLys 429
 DB 745 GAAGAAATTAATGAGAACCAAAATGGGACACCAATGATTTAGTGAATTAATCTCAAA 804
 OY 430 AsnasnValAspAsnLysAspAsnAspSer 439
 DB 805 GAGAAAGCTTTGGAATACAGATAGTCTCA 834

RESULT 11

US-10-094-240-10/C
 ; Sequence 10, Application US/10094240
 ; Publication No. US20030082637A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZWIEBEL, LAURENCE J.
 ; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
 ; FILE REFERENCE: N8289
 ; CURRENT APPLICATION NUMBER: US/10/094,240
 ; PRIORITY FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 10/056,405
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: 60/264,649
 ; PRIORITY FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 4985
 ; TYPE: DNA
 ; ORGANISM: Anopheles gambiae
 US-10-094-240-10

Alignment Scores:

Pred. No.: 0.000192 Length: 4985
 Score: 129.00 Matches: 33
 Percent Similarity: 52.38% Conservative: 22
 Best Local Similarity: 31.43% Mismatches: 32
 Query Match: 5.42% Indels: 18
 DB: 9 Gaps: 4

US-09-890-220-2 (1-445) x US-10-094-240-10 (1-4985)

OY 348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
 DB 3174 TTCTTAAGATATTAATGTTATTAATCTAAGT-----TACGGATTAATTAATTAATCT 3121
 OY 368 Ile-----LysLeuThrPAsnHisGlyLeuVal-----Asp 377
 DB 3120 ATTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3061
 OY 378 SerAlaThrIleasnasncysasnThrilleuGluasnCysArgasnSerSerVal 397
 DB 3060 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3019
 OY 398 ThrThrThrThrasnasnasnSerValAspArgProSerAspSerAsnThrAsnAsn 417
 DB 3018 AATAAGACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2959
 OY 418 AsnIleValAspHisProasnAspIleasnasnLysAsnValAspAsnLysAspAsn 437
 DB 2958 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2899
 OY 438 AsnSerArgAspLys 442

APPLICANT: Sherman, Amit
 APPLICANT: Silva, Jeff
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 FILE REFERENCE: 109272.147
 CURRENT APPLICATION NUMBER: US/09/801,368
 PRIOR FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 09/487,558
 PRIOR FILING DATE: 2000-01-19
 PRIOR APPLICATION NUMBER: US 60/160,587
 PRIOR FILING DATE: 1999-10-20
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 131
 LENGTH: 2886
 TYPE: DNA
 ORGANISM: Saccharomyces cerevisiae
 US-09-801-368-131

Alignment Scores:
 Pred. No.: 0.000118 Length: 2886
 Score: 127.50 Matches: 26
 Percent Similarity: 61.90% Conservative: 13
 Best Local Similarity: 41.27% Mismatches: 23
 Query Match: 5.36% Indels: 1
 Gaps: 1

US-09-890-220-2 (1-445) x US-09-801-368-131 (1-2886)

QY 382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
 DB 1501 AACCAACACATTAACGATTAACGATTAACATTAATTAATTAATTAACAT 1560
 QY 402 AsnAsnSerValAspArgProSerAspSerAspThrAsnAsnAsnIleValAsp 421
 DB 1561 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
 QY 422 HisProAsnAspIleAsnAsn--LysAsnAsnValAspAsnLysAspAsnSerArg 440
 DB 1621 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
 QY 441 AspLysVal 443
 DB 1681 GATTAATTAAT 1689

RESULT 14
 US-10-006-780-1
 Sequence 1, Application US/10006780
 Publication No. US20030104496A1
 GENERAL INFORMATION:
 APPLICANT: Sakowicz, Roman
 APPLICANT: Bernard, Christophe
 APPLICANT: Guo, Jun
 APPLICANT: Freedman, Richard
 TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
 FILE REFERENCE: CYTO083
 CURRENT APPLICATION NUMBER: US/10/006,780
 CURRENT FILING DATE: 2001-11-30
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 3931
 TYPE: DNA
 ORGANISM: Plasmodium falciparum
 US-10-006-780-1

Alignment Scores:
 Pred. No.: 0.000529 Length: 3931
 Score: 124.00 Matches: 96
 Percent Similarity: 36.36% Conservative: 68
 Best Local Similarity: 21.29% Mismatches: 166
 Query Match: 5.21% Indels: 122

DB: 9 Gaps: 21
 US-09-890-220-2 (1-445) x US-10-006-780-1 (1-3931)
 QY 24 IleyrCysLysProValArgLeuThrAsnIlePheHisLeuArgSerLeuGluAsnPro 43
 DB 483 ATTTATTTGGT-----AAATTAATGATTATTAACAAAGCTAGTGT----- 527
 QY 44 SerPheLeuProArgCysLeuAsnTylisIleGluValAspArgLysSerArg 63
 DB 528 -----GTACACCAAT-AGAAAAAGGCAAAAGCA 556
 QY 64 SerThrGlyMetValValPheAsnTylLysAspCysAsnSerThrLeuGluThrGlu 83
 DB 557 AGTGTAGTAAAGATTAAATTAATTAAGAGTATTAACAA-----AGAGCAATTAAT 610
 QY 84 ValArgGluAspCysSerCysProPheCysSerThrLeuGluGlySerPheGlyLe 103
 DB 611 ATTAATAATGATGATGCTTTTATTAAGAAAAATGCTTAATTAATTCACAAAGAT-- 668
 QY 103 uGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 123
 DB 669 -----GAATCATCTACATCATGCTATATTAATTAATGATTTAAAA----- 710
 QY 123 uTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePhe----- 138
 DB 711 -----GATTAATAATAATAATACATCTCTGGAAAAATTGCTTCAATGATTAGCAG 763
 QY 139 -GluGluGluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProAr 158
 DB 764 AAGTGAAGAGAGCTGAT-----ACCGTTTCACAAA 796
 QY 158 gLysArgArgGluArgGlyLysArgAsnThrArg-----ArgLeuLysValCy 175
 DB 797 TAAACAACACCAACCGATGAGCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 856
 QY 175 sPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeu 195
 DB 857 TATTCAGACTAATGATTA-----GATTAATAATCAATATACCTTCAG 898
 QY 195 uAsnAspGlyLysArgGlyLeuGlyTyrProGluAlaThrGluLeu----- 210
 DB 899 AGATTCA-----GAATTAATCAATTAAGTTTAAGCAATATAT 934
 QY 211 -AlaGlyGlnPheGlu-----MetThrSerAsnIleProAlaIleAla----- 225
 DB 935 TGTAGCAAAATCTAAAGATTAATGATGATGATTAATTTCTCAATTAATGATGATGAT 994
 QY 226 -----HisSerSerLeuAspAlaGlyAlaLysValIleLeuThr 238
 DB 995 GCAACATTTGAATCATTAATGATTTCTTCAAGAGTTAAGACCTTAATAATAATAATCTAC 1054
 QY 238 rSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGlu 258
 DB 1055 ATGCTTAATGAAGAAGATGATTAACAAATTAACGAAGATTAATTAATTAATTAATTAAT 1112
 QY 258 ArgSerHisLeuLeuLeuGluLysArgGlnPheTyrHisSerHisArgValGlnProMe 278
 DB 1113 -----GGAAGTGAATGAATGATC 1132
 QY 278 talLeuGluGluValMetSerAspArgAspSerGluAspGluValAspAspValAl 298
 DB 1133 TGTATTAAGAAATGATGAT----- 1151
 QY 298 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysG 318
 DB 1152 -----ATCAAAATCAATCATCTCTTCAAAAT-----AATTAATAATAATAATAAT 1195
 QY 318 nPheMetHisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGluHisI 338
 DB 1196 CAATGCTGCAAAATTAATGATTAAGATGAAGAAAGAAATTAATTAATTAATTAATTAAT 1255
 QY 338 eSerThrPheLysGluLysPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSe 358

Search completed: June 20, 2003, 02:35:54
Job time : 269 secs

```

DB      1256 TGATTAACCTAGGAGGATTCACATCAACCTTGGAAAA-----TACAGTTC 1303
QY      358 rLeuphetrcpCystpArgLeupheLleLysLeuptrpAsnHISGlyLeuValasp-- 377
DB      1304 ACTTAATGATATAGACAAAATTAAGAAAAATTAAG-----AAAAAGGTTAATTAATTA 1357
QY      378 -SerAlaThrIleAsnAsnCysAsnThrIleuGlu-----AsnCysArgAs 393
DB      1358 TAAAGTACATTAATATATATGACACACATAATAAAGCACACATTAATTAATTA 1417
QY      393 nSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAs 413
DB      1418 TAATAATGATATATATGATTAATTAATGATATATATATTAATTAATTAATGATAG 1477
QY      413 nThrAsnAsnAsnAsnIleValAspHis-----ProAsnAspIleAsnAsnLysAsnAs 431
DB      1478 TAGTATGTCATTAATCAATGATCAATCATATGATCAATTAATTAATTAACACAAATATCA 1537
QY      431 nValAspAsnLysAspAsnAsnSerArgasp 441
DB      1538 TCTAAATTAACACAAATTAATTAATTAAT 1568

```

RESULT 15

```

US-09-969-373-472/c
; Sequence 472, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Efferetz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 472
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-472

```

Alignment Scores:

```

Pred. No.: 7.83e-06 Length: 264
Score: 123.00 Matches: 26
Percent Similarity: 61.11% Conservative: 18
Best Local Similarity: 36.11% Mismatches: 24
Query Match: 5.17% Indels: 4
DB: 10 Gaps: 2

```

US-09-890-220-2 (1-445) x US-09-969-373-472 (1-264)

```

QY      370 LeuTPAsnHISGlyLeuValaspSerAlaThrIleAsnAsnCysAsnThrIleuGlu 389
DB      259 ATTGGAGTTC-----GTCCCAATTGACTATAAATAAAGTTTAATATCTC--- 209
QY      390 AsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValaspArgPro 409
DB      208 ---TGTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 152
QY      410 SerAspSerAsnThrAsnAsnAsnAsnIleValaspHisProAsnAspIleAsnAsnLys 429
DB      151 AATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 92
QY      430 AsnAsnValaspAsnLysAspAsnAsnSerArgasp 441
DB      91 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 56

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:44:28 ; Search time 3000 seconds
(without alignments)
3729.469 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRONCAKSPSEVISTDE.....INNNNVNDKNNRDKYIK 445

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 24791104 seqs, 12571243825 residues
Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=tlh
-Q=/cgn2_1/USPTO.spool/US09890220/runat_13062003_144853_16527/app.query.fasta_1.583
-DB=pending_Patents_NA_Main -OFMT=fastap -SUFFIX=trmp -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -ONITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09890220.ecgn.1.1.2950.etrnat_13062003_144853_16527 -NCPU=6 -ICPU=3
-NO_MAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONCLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents_NA_Main:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2378	100.0	1722	US-09-890-220-1	Sequence 1, Appl1
2	2378	100.0	1722	US-09-890-475-57	Sequence 57, Appl1
3	2363	99.4	1722	US-09-890-220-47	Sequence 47, Appl1
4	2285	96.1	1715	US-09-890-220-4	Sequence 4, Appl1
5	2285	96.1	1715	US-09-890-475-58	Sequence 58, Appl1
6	2258.5	95.0	1737	US-09-890-220-7	Sequence 7, Appl1

7	1923	80.9	1497	19	US-09-513-996A-69038	Sequence 69038, A
8	1813.5	76.3	6338	33	US-09-890-220-3	Sequence 6, App11
9	1639	69.8	5895	33	US-09-890-220-6	Sequence 587, App
10	1639	69.8	201471	31	US-09-803-859-587	Sequence 587, App
11	1639	69.8	201471	31	US-10-218-999-5373	Sequence 5373, App
12	1166	49.0	1722	42	US-09-339-947A-2	Sequence 2, App11
13	923	38.8	2280	17	US-09-303-460-1	Sequence 9, App11
14	907.5	38.2	2346	17	US-09-339-947A-9	Sequence 9, App11
15	906.5	38.1	2248	17	US-09-303-460-1	Sequence 5, App11
16	888.5	37.4	2249	74	US-09-303-460-5	Sequence 1, App11
17	871	36.6	2249	74	US-09-303-460-7	Sequence 16339, A
18	868.5	36.5	2226	74	US-09-303-460-7	Sequence 4687, App
19	833	35.0	1841	42	US-09-312-544-4687	Sequence 20106, A
20	833	35.0	1841	75	US-09-865-439A-20106	Sequence 11, App1
21	833	35.0	1856	33	US-09-303-460-13	Sequence 6, App11
22	833	35.0	2145	74	US-09-398-237-6	Sequence 26623, A
23	782.5	32.9	2296	74	US-09-654-617-26623	Sequence 3, App11
24	654	27.5	2640	17	US-09-684-016-386601	Sequence 17, App1
25	629	26.5	677	25	US-09-654-617-26623	Sequence 2870, App
26	629	26.5	677	27	US-09-684-016-386601	Sequence 1386, App
27	596.5	25.1	1176	74	US-09-303-460-3	Sequence 2869, App
28	591.5	24.9	990	25	US-09-654-617-26623	Sequence 29223, A
29	591.5	24.9	990	27	US-09-684-016-386601	Sequence 29028, A
30	559	23.5	704	74	US-09-303-460-17	Sequence 14639, A
31	527	22.2	600	74	US-10-260-238-2870	Sequence 180, App
32	527	22.2	600	61	US-09-303-460-179	Sequence 180, App
33	527	22.2	600	76	US-09-303-460-179	Sequence 180, App
34	521	21.9	492	43	US-10-260-238-1386	Sequence 180, App
35	521	21.9	492	76	US-10-260-238-1386	Sequence 180, App
36	510	21.4	726	43	US-10-260-238-1386	Sequence 180, App
37	510	21.4	726	76	US-10-260-238-1386	Sequence 180, App
38	492	20.7	2439	42	US-10-218-999-29250	Sequence 180, App
39	489	20.6	531	33	US-09-874-708A-29253	Sequence 180, App
40	489	20.6	531	65	US-09-874-708A-29253	Sequence 180, App
41	472.5	19.9	443	63	US-09-157-872-14653	Sequence 180, App
42	455	19.1	558	61	US-09-170-912-180	Sequence 180, App
43	451	19.0	453	33	US-09-890-220-13	Sequence 180, App
44	451	19.0	453	61	US-09-172-375-387	Sequence 180, App
45	443	18.6	566	33	US-09-865-439A-88732	Sequence 88732, A

ALIGNMENTS

RESULT 1
US-09-890-220-1
Sequence 1, Application US/09890220
GENERAL INFORMATION:
APPLICANT: Dean, Caroline
TITLE OF INVENTION: Methods and means for modification of plant characteristics using
FILE REFERENCE: vernalisation gene VRN2.
CURRENT APPLICATION NUMBER: US/09/890,220
PRIOR FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1722
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-1

Alignment Scores:
Score: 3.72e-230
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 33
Gaps: 0

US-09-890-220-2 (1-445) x US-09-890-220-1 (1-1722)	
QY	1 MetCysArgGlnAsnGlyAlaLysSerProGluValIleSerThrAspGlu 20
DB	231 ATGTGTAGCGCAAGATTCGCGCAAAATCCACCGAGAGATGATTTCAACTATGAG 290
QY	21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB	291 AATCTCTGAAATATTTGTAACCTGTCGATATATATACCTTTCACTTCCCTGCTGTA 350
QY	41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB	351 GGCAACCCATGTTCTTCACAGATGCTGACAAATTTGACCAAAATGGACCAAGCGCAAGA 410
QY	61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnSerThrLeuGln 80
DB	411 AAGTCAAGATCTACGGATGCGATGATTTCACTATAGATTTGATATACACTTACAG 470
QY	81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB	471 AAAATGAAAGTAAAGGAGAGATTTCTTGTGCAATTTGCTCATGCTATGCTGATGCTTC 530
QY	101 LysGlyLeuGlnPheHisLeuAsnSerHisAspLeuPheGluPheGluPheLysLeu 120
DB	531 AAGGGCGTCGAATTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 590
QY	121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB	591 TTCGAGAAATACACACAGTTAATGTTCTGTAACCTTATTCCTTCATATTTAGAGA 650
QY	141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
DB	651 GAAGGAATATGATGAGATTAATTTGACCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 710
QY	161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB	711 AGACAAAGATGTCAGAAATTAACCCAGAGAGATTAAGTAAATGCTTTTATCCGTTGAT 770
QY	181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
DB	771 TCACCCAGTTTAACATTAATGCGCAGAAATGAATCAACCTTATATATGAAACCGT 830
QY	201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB	831 GGTTTAGGATATCCCGAGCCAGACAGACCTTGGAGCAATTTGATGACCCAGCAACAT 890
QY	221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB	891 CCACCAAGCATAGCCCACTCTCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
QY	241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB	951 GGTGTGGTCCCTGCTAGACAGAAATGATTTGCTGAGGAGATCAGAGCTTGAAGC 1010
QY	261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB	1011 CACCTACTTCTTCAGAAAGCCCAATTTATCATCTTCACAGAGTCCAGCAATGGCGCTT 1070
QY	281 GluGlnValMetSerAspArgAspSerGluAspGluValAlaAspAspValAlaAspPhe 300
DB	1071 GAGCAGAGATATGTCTGACCGGATGAGGAGATGAGATGATGATGATGATGATGATGAT 1130
QY	301 GluAspArgGlnMetLeuAspAspPheValAlaValAlaAsnLysAspGluLysGlnPheMet 320
DB	1131 GAAGATTCGCGAGATGCTTGTGATGATTTGTGATGATTAATTAATTAATTAATTAATTA 1190
QY	321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspCylHisIleSerThr 340
DB	1191 CATCTTGGAACTGCTTGTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1250
QY	341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360

Db 1251 GCATGTGAACATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCATCTTTC 1310
 QY 361 TTPCYSTRPARGLEUHEULEULEYLSLEUTPRASNISGLYLEUVALSPSERALATPR 380
 Db 1311 TGGTGTGGAGATGTTTTGATTAACATGTAACACGAGCACTTGTGCATCAGCCACC 1370
 QY 381 ILEASNASCYASNTHIRILEUGLUANSCYARGASERSESPHTRHTRHTR 400
 Db 1371 ATCAACAATGCATATCCATCTCCGAGATTCGCCGTAATAGCTCAGACACACACACC 1430
 QY 401 ASNASNASNSENSEVALSPARGPROSERASPSERASNTHRASNASNSENLEVAL 420
 Db 1431 AACCAACAACAAGTGTGATGCTGCCCACTGACTCAACACCAACAATTAACATTTG 1490
 QY 421 ASPHISPROASNAPRILEASNASNLYSASNASNVALASPSNLYSPASNSENSEARG 440
 Db 1491 GATCATCCCAATGACATAAACAACAAGAAATGTTGACACAAAGACAAATTAACACAGA 1550
 QY 441 ASPLYSVALILELYS 445
 Db 1551 GACAAAGTAAATTAA 1565

RESULT 2

US-09-890-475-57
 / Sequence 57, Application US/09890475
 / GENERAL INFORMATION:
 / APPLICANT: Johanson, Urban
 / APPLICANT: West, Joanne
 / APPLICANT: Dean, Caroline
 / TITLE OF INVENTION: Arabidopsis thaliana derived Frig1da gene conferring late flowering
 / FILE REFERENCE: Newburn
 / CURRENT APPLICATION NUMBER: us/09/890,475
 / CURRENT FILING DATE: 2001-11-13
 / PRIOR APPLICATION NUMBER: PCT/GB00/00197
 / PRIOR FILING DATE: 2000-01-25
 / PRIOR APPLICATION NUMBER: GB 9902660.1
 / PRIOR FILING DATE: 1999-02-05
 / NUMBER OF SEQ ID NOS: 58
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 57
 / LENGTH: 1722
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Description of Artificial Sequence: Landsberg
 / US-09-890-475-57

Alignment Scores:

Pred. No.: 3,72e-230 Length: 1722
 Score: 2378.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 33 Gaps: 0

US-09-890-220-2 (1-445) x US-09-890-475-57 (1-1722)

QY 1 METCYSTRPARGLEUHEULEULEYLSLEUTPRASNISGLYLEUVALSPSERALATPR 20
 Db 231 ATGTGTAGGACAGATTTGTTGATTAACATGTAACACGAGCACTTGTGCATCAGCCACC 290
 QY 21 ASNLEULEULEYCYSLYSPROVALARGLEUTPRASNILEPHEHISTLEUARGSERLEU 40
 Db 291 AATCTTTATATATTGTAACCTGTTCCACTATATATACCTTTTACCTTGCCTCTCTA 350
 QY 41 GLYSNPROSEPHLEUPROARGCYSEUASNTRYLYSILEGLYALALYSARGLYSARG 60
 Db 351 GGCAACCAATCGTTTCTTCCAGAGTCTGAAGTGAATAAATTGGACAAAGCCAAAGAGA 410
 QY 61 LYSERARGSERTHIRIGLYEVALVALPHEASNTYLYSPCYASNASNTHIRLEUGIN 80
 Db 411 AAGTCAAGATCTACTGGATGTAGTTTCACTAATAAAGGATTTGTAATAACACATTAACAG 470

QY 81 LYSTHRGLUVALARGLUASPCYSSERCYSPPROPHECYSEMERLEUCYGLYSERPE 100
 Db 471 AAACTGAGCTTAGGGAGATTTGTTCTTCATTTTGTCTATGCTATAGTGTGACTTC 530
 QY 101 LYSGLYLEUGLNPHEHISLEUASNSENSEHISASPLEUPHEGLUPHEGLUPHELYLEU 120
 Db 531 AAGGGGCTCAATTTTCATTTGAATTCATCATGATTTATTTGAATTTGATTTCAAGCTT 590
 QY 121 PHEGLUGLTYRGLINTHRVALSVALSERVALYSLEUASNSEPHLEPHEGLUGLU 140
 Db 591 TTCGAAGAATTACAGACAGCTTAATGTTTGTAAAACCTTAATCCCTTCATATTAGAGA 650
 QY 141 GLUGLISERASPSAPRILYSPHEGLUPROPHESERLEUCYSESLYSPROARGLYSARG 160
 Db 651 GANAGAGATGATGACATTAATTAATTTGAGCCCTTCTCTCTGCTCGAACCCTGTAAAGCG 710
 QY 161 ARGGLNARGGLYGLYARGASNSENTHRARGARGLEUVALCYSPHEUPROLEUASP 180
 Db 711 AGACAAGAGAGTGGCAGAAATATAACACAGAGACTTAAGTATGCTTTTACCGTTGGAT 770
 QY 181 SERPROSERLEUTHRASNGLYTHRGUASNGLYLETHIRLEULEUASNAPGLYASNARG 200
 Db 771 TCACCCAGTTTAATGATGACAGCAAGAAATGGAATCCCTTAATATGGAACACCGT 830
 QY 201 GLYLEUGLYTPROGLUALATHRGULEUALAGLGLNPHEGLUMETHRSEARSNILE 220
 Db 831 GGTTTAGATATCCGAGGCAACAGACTTCTGCGCAATTTGAGATGACCAACACATTT 890
 QY 221 PROFOALALEALHISSESERLEUASPARALAGLYALALYVALILEUTHRSEGLU 240
 Db 891 CCACACGACATAGCCCACTCTCTGAGCGGTGGCTAAAGTATATGACAAAGCGAA 950
 QY 241 ALAVALVALPROALATHRYSTRHARGLYSLEUSERALAGLUARGSERGLUALARGSER 260
 Db 951 GCTGTGTCCTCCCTACTAAGACAAAGAAATATCTGCGACGACGACAAAGGCTAAGAC 1010
 QY 261 HISLEULEULEUGLINSARGLNPHEUTYHISSENHISARGVALGINPROMETALAU 280
 Db 1011 CACCTACTCTTCAGAAAGCCCAATTCATCTTCACAGAGTCCAGCAATGGCGCTT 1070
 QY 281 GLUGLINALMETSERASPARARGSPSERGLUASPGULVALASPSAPRYVALALASPHE 300
 Db 1071 GAGCAAGTAAATGCTGACCGGAGTACGAGATGAAGTGCATGACATGTGCACAGATTTT 1130
 QY 301 GLUASPARGLIMETLEUASPPRHEVALASRYVALASNLYSASPGULYSGINPRHEMET 320
 Db 1131 GAAGATCGCCAGATGCTTGATGACTTTGTGATGTGAATTAAGATGAAGCAAAATTCATG 1190
 QY 321 HISLEUTPRASNSEPHHEVALARGLYGINARGVALILEALASPGLYHISILESETRP 340
 Db 1191 CATCTTTGGAACCTGTTTGTAGAAAACAAGGTTATAGCAGATGTGATATCTTGTGG 1250
 QY 341 ALACYSGLUALAPHESEARGPRHEUTYGLYUGLYLEUHNISARGTYRSESERLEUPE 360
 Db 1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTATCATCATCTTTC 1310
 QY 361 TTPCYSTRPARGLEUHEULEULEYLSLEUTPRASNISGLYLEUVALSPSERALATPR 380
 Db 1311 TGGTGTGGAGATGTTTTGATTAACATGTAACACGAGCACTTGTGCATCAGCCACC 1370
 QY 381 ILEASNASCYASNTHIRILEUGLUANSCYARGASERSESPHTRHTRHTR 400
 Db 1371 ATCAACAATGCATATCCATCTCCGAGATTCGCCGTAATAGCTCAGACACACACACC 1430
 QY 401 ASNASNASNSENSEVALSPARGPROSERASPSERASNTHRASNASNSENLEVAL 420
 Db 1431 AACCAACAACAAGTGTGATGCTGCCCACTGACTCAACACCAACAATTAACATTTG 1490
 QY 421 ASPHISPROASNAPRILEASNASNLYSASNASNVALASPSNLYSPASNSENSEARG 440
 Db 1491 GATCATCCCAATGACATAAACAACAAGAAATGTTGACACAAAGACAAATTAACACAGA 1550

OY 441 Asplysvalielys 445
Db 1551 GACAAAGATTAATTA 1565

RESULT 3

US-09-890-220-47
Sequence 47, Application US/09890220

GENERAL INFORMATION:

APPLICANT: Dean, Caroline
APPLICANT: Gendall, Anthony
TITLE OF INVENTION: Methods and means for modification of plant characteristics using
FILE REFERENCE: Mewburn
CURRENT FILING DATE: 2001-07-27
CURRENT FILING DATE: 2001-07-27
PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
PRIOR FILING DATE: 1999-01-28
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 47
LENGTH: 1722
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-47

Alignment Scores:

Pred. No.: 1,24e-228 Length: 1722
Score: 2363.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.37% Indels: 0
DB: 33 Gaps: 0

US-09-890-220-2 (1-445) x US-09-890-220-47 (1-1722)

OY 1 Metcysarglnasncysatgatalyserserprogluvaliieserthrassp 20
Db 231 ATGTGAGGCGAATTTGCGCGAATTCCTCCGCGAGCAAGTGTTCACACTATAG 290
OY 21 Asnleuileuiletyrcyslyspovalargleutyasnilephehisleuargser 40
Db 291 AATCTCTGATATATTTGTAACCTGTTCCACTATATATACATCTTCACCTGCTCTA 350
OY 41 Glysarpserpheleuproargcysleuasntyllysileglalalysarglysarg 60
Db 351 GCGAACCCATCGTTCTCCAGATGCTGAACTACAAAATTGAGCAAAAGCGCAAAAGA 410
OY 61 LysSerArgSerThrGlyMetValIlePheAsnTyrLysAspCysAsnAsnThrLeuGln 80
Db 411 AAGTCAGATCTACTGAGTGTGATGTTTCACTATATAGCATTTGTAATACATTCAG 470
OY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlnSerPhe 100
Db 471 AAAACGTAGTTAGGAGGAGTGTCTGTGCAATTTGCTCATATGTCATGCTTC 530
OY 101 LysGlyLeuGlnPheHisLeuAsnSerHisAspLeuPheGluPheGluPheLysLeu 120
Db 531 AAGGCGCTGCAATTTCTTTGATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590
OY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGln 140
Db 591 TTCCAGAAATACCGACGAGTTAATGTTCTGTAATAACTTAATCTTCATATTTGAGAA 650
OY 141 GluGlySerAspAspLysPheGluProPheSerLeuGlnSerLysSerLysProArgLysArg 160
Db 651 GAAAGAGATGATGCGATAATTTGAGCCCTTCTCTGCTCCAAAACCTCGTAAGCGG 710
OY 161 ArgGlnArgLysGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
Db 711 AGACAAAGAGGTGCGAATAATACACAGAGACTTAAGATGATCTTTTATACCGTTGAT 770

OY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuAsnAspGlyAsnArg 200
Db 771 TCACCGAGTTTAACATATGCGACAGAAATGATATCACTACTTAATATGTAACACGT 830
OY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGlnMetThrSerAsnIle 220
Db 831 GGTTRAGATATCCCGAGCGAACAGACTTCTGCAATTTAGATGACGACACAT 890
OY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
Db 891 CCACCGACCATGCGCCACTCTCTCTGACCTGTGCTTAATGTAATGACACCGAA 950
OY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
Db 951 GCTGTGTCCTCGTACTAGACAGAAAGTATCTGCTGAGCATCAGAGCTAGAACG 1010
OY 261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
Db 1011 CACCTACTCTTCAGAAAGCGCAATTCATCTCATCAGAGCTCCAGCCAAATGCGCTT 1070
OY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
Db 1071 GAGCAAGTAAATGCTGACCGGATAGCGAGATGAAGTCATGACATGCTTGCAGATTT 1130
OY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
Db 1131 GAAGATCGCCACATGCTTGATGACTTGTGTGATGTAATTAAGATGAAGCAATTCATG 1190
OY 321 HisLeuTyrPheSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
Db 1191 CATCTTGGAATCGCTTTGTAAGAAACAAAGGCTTTACCAATGCTATCTCTTGG 1250
OY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
Db 1251 GCATGTGACGATTTTTCAGATTTTACAGAAAGATGTCACCGCTTTCATCATCTTTC 1310
OY 361 TrpCysTyrPheGluPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
Db 1311 TGGTGTGAGATTTGTTTATTAATTAACATAGAAACCATGACCTGCACTCAGACGAC 1370
OY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
Db 1371 ATCAACAACATGCAATCCATCTCGAGAAATGCGGTAAATGCTCAACACACCCAC 1430
OY 401 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
Db 1431 AACCAACAACAAGTGTGATGTCGCCAGTACCAACCAACCAATTAACATTTGTG 1490
OY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValIleAspAsnLysAspAsnSerArg 440
Db 1491 GATCATCCCAATGACATTAACACACAGAACAAATGTTACAAACAGGACATTAACAC 1550
OY 441 Asplysvalielys 445
Db 1551 GACAAAGATTAATTA 1565

RESULT 4

US-09-890-220-4
Sequence 4, Application US/09890220

GENERAL INFORMATION:

APPLICANT: Dean, Caroline
APPLICANT: Gendall, Anthony
TITLE OF INVENTION: Methods and means for modification of plant characteristics using
FILE REFERENCE: Mewburn
CURRENT FILING DATE: 2001-07-27
CURRENT FILING DATE: 2001-07-27
PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
PRIOR FILING DATE: 1999-01-28
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 4
 LENGTH: 1715
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-890-220-4

Alignment Scores:

Pred. No.:	1,01e-220	Length:	1715
Score:	2285.00	Matches:	428
Percent Similarity:	97.518	Conservative:	3
Best Local Similarity:	96.838	Mismatches:	9
Query Match:	96.09%	Indels:	2
DB:	33	Gaps:	1

US-09-890-220-2 (1-445) x US-09-890-220-4 (1-1715)

```

QY      1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB      233 ATGTGTAGGCAAAATTTGCGCCGAAATCCTCACCGAGGAGATGATTTCAACCTGATGAG 292
QY      21 AsnLeuLeuIleTyrCysAlaProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB      233 ATGTGTAGGCAAAATTTGCGCCGAAATCCTCACCGAGGAGATGATTTCAACCTGATGAG 292
QY      23 AATCTCTGATATATTGTAACCTTGTGACTATATTAACATCTTTCACCTTGCTCTCTA 352
DB      41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
QY      353 GGCAACCATGCTGTTCTGCAAGATGCTTGACACTCAAAATTTGGGGCAAAAGCGCAAAAGA 412
DB      61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
QY      413 AAGTCMAAGATCTACTGAGATGCTGACTTTTCAACTTAAAGGATTTAATATACATTACNA 472
DB      81 LysThrGluValArgGluAspCysSerCysPropheCysSerMetLeuGlySerPhe 100
QY      473 AGAATCTGAAGTATAGGAGAGATTTGTTCTTGTCCATTTTGTCTATATGCTGATGCTTC 532
DB      101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
QY      533 AAGGGGCTGCATATTCATTGAAATTCATCCTCATGATTTATTTGAATTTGAGTTCAAGCT 592
DB      121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
QY      593 TTGGAAGATACCAACAGATTAATGTTCTGTAAACATTAATTCCTTCATATTTAGAGNA 652
DB      141 GluGlySerAspAspAspLysPheGluPheSerLeuPheSerLeuPheSerLeuPheSerLeu 160
QY      653 GAAGAAAGTATGATGATTAATTTGAGCCCTTCTCTCTGCTCGAAGCTGTAAGCT 712
DB      161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
QY      713 AGACAAAGAGGTGGAGAAATTAACACAGAGACTTAAAGTATGCTTTTACCGTTGAT 772
DB      181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
QY      773 TCACCCAGTTTAACTAATGCGACAGAAATGGAATTCCTGCTGATGATGTAAGCCGT 832
DB      201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluIleMetThrSerAsnIle 220
QY      833 GGTTTAGGATATCCCGAGGCAACAGAGCTGCTGACAAATTTGAGATGATGCTGACAACT 892
DB      221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
QY      893 CCACAGGCATAGCCACCTCTCTGACGCTGGTGCTAAAGTTATATATTAACAACCGNA 952
DB      241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
QY      953 GCTGTGCTCCCTGCTACTAAGCAAGAAAGTATCTGCTGACGATCAAGGCTGGAAC 1012
DB      261 HisLeuLeuGluLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
QY      1013 CACCTACTTCTTCAAGAAACGCAATTTCTATCTCTCACAGAGTCCAGCAATGGCGCTT 1072
DB      281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
  
```

```

DB      1073 GAGCAAGTAATGTCGATGGGATAGCGAGATGAAGTGCATGACATGTCACATTTT 1132
QY      301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB      1133 GAAGATCCGCAAGATCTTGATGATCTTGTGGATGATGAATGAAGATGAAGCAATTCATG 1192
QY      321 HisLeuTyrPheSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB      1193 CATCTTTGACATCTGTTTGTATGAAGAAACAAAGGTTATAGCAGATGTCATATCTCTGG 1252
QY      341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
DB      1253 GCATGGAAGTATTTTCAAGATTTTACGAGAAAGAGTTGCACTGTACTCATCACTCTTC 1312
QY      361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrPheAsnHisGlyLeuValAspSerAlaThr 380
DB      1313 TGGTGTGGAGATTTGTTTGTATTAACATATGAAACATGAGACTGTGACTCAGCCACC 1372
QY      381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
DB      1373 ATCAACACTGCAATATCCATCTCGAAGATTCCTGATATCTCA-----GTCACTAAC 1426
QY      401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
DB      1427 AACCAACACACAGATGTGATCATCCAGTCACTCAACACCAACACAAATTAATTTGTG 1486
QY      421 AspHisProSerAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB      1487 GATCATCCGAAATGACATATAAAACAAAGAAATGTTGACMACAGACAAATTAACAGAGA 1546
QY      441 AspLys 442
DB      1547 GACAAAG 1552

RESULT 5
US-09-890-475-58
; Sequence 58, Application US/09890475
; GENERAL INFORMATION:
; APPLICANT: Johanson, Urban
; APPLICANT: West, Joanne
; TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late flow
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/890,475
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/GB00/00197
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: GB 9902660.1
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Columbia VRN2
US-09-890-475-58

Alignment Scores:
Pred. No.: 1,01e-220 Length: 1715
Score: 2285.00 Matches: 428
Percent Similarity: 97.518 Conservative: 3
Best Local Similarity: 96.838 Mismatches: 9
Query Match: 96.09% Indels: 2
DB: 33 Gaps: 1

US-09-890-220-2 (1-445) x US-09-890-475-58 (1-1715)

QY      1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
  
```

Db	233	ATGTTGATGAGCAAAATTTCTGGCCGCGAAATCCTCACCGGAGGAAGTATTCTCAACGATGAG	292
OY	21	AsnIeuLeuIleTyrCysLysProValArgLeuTyrAsnIleIlePheIleIleuAspSerLeu	40
Db	293	AATCTCTGATATATTGTAACTTGAACCTGTGTGACATATATACTCTTCACTTCGCTCTCTA	352
OY	41	GLYAsPProSerPheLeuProArgCysLeuSerTyrLysIleGlyAlaLysArgLysArg	60
Db	353	GGCAACCCATCGTTTCTGCCAAGATGCTTGAACTCAAAATATGGGCGAAAGCGAAAGA	412
OY	61	LysSerArgSerThrCysLeuValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln	80
Db	413	AAGTCAGATCTACTGGAGTGGTACGTTTCACTTAAGATTTGATTAATATACATTTACAA	472
OY	81	LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe	100
Db	473	AGAACCTGAATTTAGGGAGGATGTTCTGTCTCATTTTCTCTTAAGCATGTGTGGTACGTC	532
OY	101	LysGlyLeuGlnPheIleLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu	120
Db	533	AAAGGGCGTGCATTTCTATTGGAATTCATCATCATGATTTATTTGGAATTTGAGTTCAAGCTT	592
OY	121	PheGluGluTyrGlnThrValAsnAlaSerValLysLeuAsnSerPheIlePheGluGlu	140
Db	593	TTGGAAAGAATACCACGACAGTATATGTTCTGTGAAAACATTAATTCCTTCATATTGGAGAA	652
OY	141	GluGluSerAspAspAspLysPheGluProPheSerLeuCysSerLysProAlaGlyAsp	160
Db	653	GAAGGAAGTATGATGATTAATTTGAGCCCTTCCTCTCTGCTGCAAACTCGTAAAGCGT	712
OY	161	ArgGluArgGluValArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp	180
Db	713	AGACAAAGAGGTGGCGAAATACACACGAGAGCTTAAGATGCTTTTACCGTTGGAT	772
OY	181	SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg	200
Db	773	TCACCCAGTTTACGTAATATGGCACAAATAATGCAATTCGCCCTCGATATATGGAAACCGT	832
OY	201	GlyLeuGluTyrProGluValaThrCysLeuAlaGlyGlnPheGluMetThrSerAsnIle	220
Db	833	GCTTTAGCATTCGCCAGGACACACAGCTTCTGTGCACATTTTGATGAGACTACACATTT	892
OY	221	ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu	240
Db	893	CCACCAACCATACCCACTCTTCTGTGACCTGTGTCTTAAGTATATTAAACACCGAA	952
OY	241	AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer	260
Db	953	GCTGTGTGCTCCGCTACTACAGCAAGAAATTAATCTCTGACGATCAGAGGTAAAGC	1012
OY	261	HisLeuLeuLeuGluLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu	280
Db	1013	CACATCACTTCTCAAAAGCCCAATTCATATCTCACAGAGTCCAGCCAAATGGCGCTT	1072
OY	281	GluGlnValMetSerAspArgAspSerGluAspLysAlaLysAspAspValAlaAspPhe	300
Db	1073	GAGCAAGTATGTCTGATCGGATCGGATAGCGAGGATGAATCATACATACGATGTCAGATTTT	1132
OY	301	GluAspArgGluMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet	320
Db	1133	GAAGATCGCCCAATCTGTGATGACTTTTGTGATGTGAATAAGTGAAGAAAGCAATTCAG	1192
OY	321	HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr	340
Db	1193	CATCTTTGACATCGTTTGTAAAGAAACAAAGGTTTATACAGATGCTCATATCTCTTG	1252
OY	341	AlaCysGluAlaPheSerArgPheTyrGluLysLeuMetHisArgTyrSerSerLeuPhe	360
Db	1253	GCATGTGAAGTATTTTCAAGATTTTACAGAAAGAGTTCGACGCTTATCATTACACTCTTC	1312
OY	361	ThrProTyrArgLeuPheLeuIleLysLeuTyrAsnHisGlyLeuValAspSerAlaThr	380
Db	1313	TGGTGTGGAGAGTGTTTTATTAATATAGGAACCATGACCTGTGTGACATCAGCCAC	1372

Accession	Gene	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
OY 381	ILASASASACysmsnthrilleleuGluasnCysArGmsnserserSptrhrThrThr	2258.50	1737	428	3	9	10	2
Db 1373	ATCAAACTGCATACATCTCTGAGAAATGGCCGTAAATACATCA-----GTCATAC	95.78%						
OY 401	ASASASASASerValAspAprProSerSerserSnrhrsnmsnslleVal	2258.50	1737	428	3	9	10	2
Db 1427	AAACAACAACAGTGTGATCATCCAGTACCTCAACCAACAACATATACATG	95.78%						
OY 421	ASHisProAsnAspIleAsnAsnlysnmsnslValAspAsnlysnAspAsnserATG	2258.50	1737	428	3	9	10	2
Db 1487	GATCATCCCAATGACATATAAACAAGACAAATGTGACAAACAAGACATACACAGA	95.78%						
OY 441	AsPlys 442	2258.50	1737	428	3	9	10	2
Db 1547	GACCAAG 1552	95.78%						
<p>US-09-890-220-7</p> <p>Sequence 7, Application US/09890220</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Dean, Caroline</p> <p>APPLICANT: Genedall, Anthony</p> <p>TITLE OF INVENTION: Methods and means for modification of plant characteristics us</p> <p>FILE REFERENCE: Newbury</p> <p>CURRENT APPLICATION NUMBER: US/09/890, 220</p> <p>PRIOR FILING DATE: 2000-01-28</p> <p>PRIOR APPLICATION NUMBER: GB 9901927.5</p> <p>PRIOR FILING DATE: 1999-01-28</p> <p>NUMBER OF SEQ ID NOS: 77</p> <p>SOFTWARE: PatentIn Ver. 2.1</p> <p>SEQ ID NO 7</p> <p>LENGTH: 1737</p> <p>TYPE: DNA</p> <p>ORGANISM: Arabidopsis thaliana</p> <p>US-09-890-220-7</p>								
<p>Alignment Scores:</p> <p>Pred. No.: 5e-218</p> <p>Score: 2258.50</p> <p>Percent Similarity: 95.78%</p> <p>Best Local Similarity: 95.11%</p> <p>Query Match: 94.97%</p>								
US-09-890-220-2 (1-445) x US-09-890-220-7 (1-1737)								
OY 1	MetCysArGlnAsrCysArGAlaLysSerProGluValIleSerThrAspGlu	2258.50	1737	428	3	9	10	2
Db 233	ATGTGAGGCAAAATGTCGCGCAAAATCCACCCGGAGAAATGATTCATCACTAGAG	95.78%						
OY 21	AsnLeuLeuIleTyCysLysProValArgLeuTyAsnIlePheHisLeuArgSerLeu	2258.50	1737	428	3	9	10	2
Db 293	AATCTCTGATATATATGTAACCGTGTGCACTAATAAATCTTTTACCTTGCGCTCTA	95.78%						
OY 41	GlyAsnProSerPheLeuProArgCysLeuAsnTyrlLysIleGlyAlaLysArGlyAsrG	2258.50	1737	428	3	9	10	2
Db 353	GCGCAACCATCTCTTCTGCCAAGATGCTTACACTACAAATTTGGGCAAGCGCAAGA	95.78%						
OY 61	LysSerArgSerThrGlyMetValValPheAsnTyrlLysAspCysAsnAsnThrLeuGln	2258.50	1737	428	3	9	10	2
Db 413	AGTCAAGATCTACAGCGGATGCTTCACTATATAGAGATTTGTAATATACATACAA	95.78%						
OY 81	LysThrGluValArgGluAspCysSerCysProPheCysSerMetIleuCysGlySerPhe	2258.50	1737	428	3	9	10	2
Db 473	AATCTGATGATGAGAGATGTTCTTGCCATTTTGCTATGCTATGATGAGTACCTTC	95.78%						
OY 101	Lys-----GlyLeuGlnPheHisLeuAsnSerSerHisAspLe	2258.50	1737	428	3	9	10	2
Db 533	AAGGGGGCACTATTACAACTGAGGGCGCAATTTCAATTGGAATTCATCTCAAGATT	95.78%						

QY	113	uphegluphgeluphelystleuphegluglulutyglulnthValasnaIservalysle	133
Db	593	ATTTGAATTTGAGTTCAAGCTTTTGGAAAGATTACACAGTTAATTTCTTGTGAACACT	652
QY	133	uasnSerPhelIephegluglulugllySerAspAspAspIyspPhegluProPheSerIe	153
Db	653	TAATTCCTCATATTGAGGAAGAAGAGATGATGATTAATTTAGACCTTCTCTCT	712
QY	153	wCySSerLysProArGLysArGaRglnArGlyGlyArGAsnaSnThrArGLeuLy	173
Db	713	CTGCTCCAAACCTTCGTAAAGGTATACAAAGAGTGGCAGAAATTAACCCAGAGAACTTAA	772
QY	173	sValCySPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleTh	193
Db	773	AGTATGCTTTTACCCTGGATTGATTCACCAGTTTACGTAAATGGCCACAAATATGGAATTC	832
QY	193	rLeuLeuAsnaSpIySnaRgIyLeuGlyTyRProGluValaThrGluLeuAlaGlyI	213
Db	833	CCCTGCTGAAGATGGAAACCGTGTTTATAGATATCCGAGGCAACAGAGCTGCTGGACA	892
QY	213	nPhegluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAl	233
Db	893	ATTTGAGATGACCTGTGCAACATTCACACAGGCATGCCACTCTTCTGTGACCGCTGTGC	952
QY	233	aLysValIleLeuThrSerGluValaValaProAlaThrLysThrArGLysLeuSerAl	253
Db	953	TAAAGTATATATTAAACAACCGAAGGTGGTGTCCTGCTACTTAAGCAAGAAATGATGTC	1012
QY	253	agLuarSerGluValaRgSerHisLeuLeuGlnLysArGlnPheTyHisSerH	273
Db	1013	TGACCGATACAGAGCTGTGAAGCCACACTTCTTCAGAAAGCCAACTTATCATCTGTCA	1072
QY	273	sArGValGlnProMetAlaLeuGluGlnValMetSerAspArGAspSerGluAspGluVa	293
Db	1073	CAGAGTCCAGCAATGCGCTGTGCAAGTATGTCATCGAGATGCGAGGTGAGT	1132
QY	293	lAspAspAspValAlaAspPheGluAspArGlnMetLeuAspAspPheValaAspValAs	313
Db	1133	CGAGGAGAGTTCACAGATTTTGAAGATCCGCAACATGCTTGATCATCTTGGATGTGAA	1192
QY	313	nLysAspGluLysGlnPheMetHisLeuTyPAsnSerPheValArGLysGlnArGValIl	333
Db	1193	TAAAGATGAAGAAGCAATTCATGCATCTTTGGAACTCGTTGTAAAGAAACAAAGGTTAT	1252
QY	333	eAlaAspGlyHisIleSerThrAlaCysGluAlaPheSerArGheTyRGluySGluLe	353
Db	1253	AGCAGATGGTATATTCCTTGGGATGGAAGTATTTTCAAGATTTTACGGAAGAAGTT	1312
QY	353	uHlArGtyrSerSerLeuPheTyRProCysTyRArGLeuPheLeuIleLysLeuTyRAsnHl	373
Db	1313	GCACCTGTATTCATCACCCTCTCTGCTGTGGAGATGTTGTTTGTATTAACATGGAACA	1372
QY	373	sGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArGAs	393
Db	1373	TGGACTTGTGCACACACCAACATCAACAACTGCAATACCATCTCCAGAAATTCGCTTAA	1432
QY	393	nSerSerAspThrThrThrAsnaAsnaAsnSerValAspArGProSerAspSerAs	413
Db	1433	TACCTCA-----GTCACTAACAAACAACAACAGGTGGATATCTCCATGACTCAAAA	1486
QY	413	nThrAsnaAsnaAsnIleValaAspHisProCAsnAspIleAsnAsnLysAsnaAsnValAs	433
Db	1487	CACCAACAACAATAACTTGTGGATTCATCCGAATAGACATATAAAACAAGAACAATGTGTA	1546
QY	433	pAsnLysAspAsnaSnSerArGAspLys 442	
Db	1547	CAACAAGGACAAATTAACAGACAGAGACAAG 1574	

QY 260 -----SerHisLeuLeuGlnLysArgGlnPheTyrHis 271
 Db 913 ATATTAAACCGAATAATTTCTTCACCCACCTTCTTCACAAACGCCAATTCATAT 972
 QY 272 SerHisArgValGlnProMetAlaLeuGlnValMetSerAspArgAspSerGlnAsp 291
 Db 973 TCTCAGAGAGTCCGCAATGGCGCTTGAGCAAGTATCTGACCGGATGCGAGAT 1032
 QY 292 GluValAspAspValAlaAspPheGluAspArgGlnMetLeuAspPheValAsp 311
 Db 1033 GAAGTCGATGACGATGTTGCGATTTGAAGATCGGACAGATCTGATGACTTTGTGAT 1092
 QY 312 ValAsnLysAspGlnLysGlnPheMetHisLeuTyrPheSerPheValArgLysGlnArg 331
 Db 1093 GTGAATTAAGATGAAAGCAATTCATGACTCTTGGAAGCTGTTGAAGAAACAAAGG 1152
 QY 332 ValIleAlaAspGlnHisIleSerTyrPheAlaCysGluAlaPheSerArgPheTyrGlnLys 351
 Db 1153 GTTATACGACAGATGCTATTTCTGCGGATGGAAGCATTTTCAAGATTTTACAGAAA 1212
 QY 352 GluLeuHisArgTyrSerSerLeuPheTyrPheTyrPheGlnPheLeuIleLysLeuTyr 371
 Db 1213 GAGTTGACCGCTTACATCATCTCTGCTGTTGAGATGTTGTTGATTAACTATGCG 1272
 QY 372 AsnHisGlyLeuValAspSerAlaThrIleAsnAspCysAsnThrIleGlnGlnAsnGly 391
 Db 1273 AACCATGACTGTGCGACTGACGCTACGACCATCACTCACTCACTCACTCACTCACT 1332
 QY 392 ArgAsnSerSerAspThrThrThrAsnAsnAsnAsnSerValAspArgProSerAsp 411
 Db 1333 CGTATAGCTTCAGACACCCACCCACCAACAAACAGAGTGTGATGCTCCAGTAC 1392
 QY 412 SerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsn 431
 Db 1393 TCAAACACCAACCAATTAACATGAGTCACTCCCAATGACATGACATCAACAAACAA 1452
 QY 432 ValAspAsnLysAspAsnAsnSerArgAspLysValIleLys 445
 Db 1453 GTTGACAAACAGGACATTAACAGCAAGCAACAAAGTAATTTAA 1494
 RESULT 8
 US-09-890-220-3
 ; Sequence 3, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE OF INVENTION: vernalisation gene VRN2.
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; PRIOR FILING DATE: 1999-01-28
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 6338
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-3

Alignment Scores:

Pred. No.: 4,2e-172 Length: 6338
 Score: 1813.50 Matches: 440
 Percent Similarity: 43.58% Conservative: 1
 Best Local Similarity: 43.48% Mismatches: 4
 Query Match: 76.26% Indels: 569
 DB: 33 Gaps: 13

US-09-890-220-2 (1-445) x US-09-890-220-3 (1-6338)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGlnIleValIleSerThrAspGln 20
 Db 2342 ATGTGAGGACCAAAATTTGTCGGCAAAATCCACCGGAGGAAATGATTTCACTGATGAG 2401
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 Db 2402 AATCTCTGATATATATGTAACCTGTCGATATATATACATCTTTCACCTTCCTCTCT 2461
 QY 41 GlyAsn----- 42
 Db 2462 GCCAAGCATATTTGCCCTCCTCCTCATCATATATAGTCAGTAATCTTTCATCTCTG 2521
 QY 42 ----- 42
 Db 2522 TGTAGATCACCACTAATAGTTTGTGATTTGCTAAGCTGATATAGTCTGATTCAGCGA 2581
 QY 43 -----ProSerPhe 45
 Db 2582 GATGTGCTCTTTTGTCTCTTAATTTGAACCTGTGTTGTTGTTGCGACGACATCTT 2641
 QY 46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer----- 62
 Db 2642 CTTCGAGATGCTTTAATCTAACAATTTGAGCAAAAGCGCAAAAGAAAGTA-TGTTTCTT 2700
 QY 62 ----- 62
 Db 2701 CTTGAATGAGCTGCTACAGTGAATGTTATTTATCTTACTTCTAATATGAAAGCTGATG 2760
 QY 62 ----- 62
 Db 2761 ACCATATTATCTTGTGAGTACATATGACATATGAAAGTTTCTTCTTGTTCATG 2820
 QY 63 -----ArgSerThrIleVal 68
 Db 2821 CTAATACTTACATTTATATAAATGTTGTTGTTAGTCAAGTCACTCTGGAATGTA 2880
 QY 69 ValIleAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 Db 2881 GTTTTCAACTATTAAGATTTAATTAACATTAACAGAAACGTAAGTATGTTTCT 2940
 QY 83 ----- 83
 Db 2941 GTTCCTGACAAAATTCGATGCTCAATGCTATGTTCTAGATGATTTGTTATTACTA 3000
 QY 84 -----ValArgGluAspCysSerCysProPheCysSerMetLe 96
 Db 3001 TTTTCTGATTTGTCATGATGAGTATGAGGAGATGTTGTTGCTATGCTATGCT 3060
 QY 96 uCysGlySerPheLys----- 101
 Db 3061 AGTGTGATGCTTCAAGGAGGACACTATTAACACTGAGGTTCTCCGGGGCTTTCATA 3120
 QY 101 ----- 101
 Db 3121 TCTAAGCTGGAATGCTACTGCTGTTTCACTGCTGATTAATCTTCACTGTTGTTACT 3180
 QY 101 ----- 101
 Db 3181 ATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3240
 QY 102 -----GlyLeuG 104
 Db 3241 GAGAAACATGCTCCACATGAGCTTCAACCAATTTGCTGTGCTATGACAGGAGCTGC 3300
 QY 104 IAPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu----- 120
 Db 3301 AATTTCAATTTGAATTCATCATCATGATTTATTTGAATTTGAGTTCAAGGATGATGTT 3360
 QY 120 ----- 120
 Db 3361 TCGAAATCTGATTTGCTATGCTTATTAATGAGTTATAGTTAAAAAGGCTCTTTC 3420
 QY 121 -----Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136

Db 3421 CTAATGAGCTTTGGAGAAATACAGACAGTAATGTTCTGTAATAACTTAATTCCTTC 3480
 Oy 137 IIEPHEGL----- 139
 Db 3481 AATATGAGGTCAGTACTTAATTAAGTGGTAATCCAGATAGCTGTGAAT 3540
 Oy 140 -----GluGluGlySerAspAspLysPheG1 149
 Db 3541 TTTGTTAATTCATCCTTATTGTTAGTACAGAAAGAAAGTATGATGAGATTAATTTGA 3600
 Oy 149 uProPheSerLeu----- 153
 Db 3601 GCCCTCTCTCTCTGTAACCTCAGAACCCCTCGATTAATACCTTAATAGCAGTAAC 3660
 Oy 154 -----CysSerly 156
 Db 3661 TCCCTGCTCTCTGTCAGTACATCTCTGTAATCCAACCAATTAATGTTTGGACGCTCGAA 3720
 Oy 156 sProAlaGlySarGArgLInArgGlyLysArgAsnThrArgArgLeuLysValCysPhe 176
 Db 3721 ACCGTGTAAGCGGAGACAAAGAGGTGGAGAAATAACACAGAGACTTAAGTATGCTT 3780
 Oy 176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyLeuThrLeuLeuAs 196
 Db 3781 TTTACCGTTGGATCACCCAGTTTAACTAATGACAGAAATGGAATCACCCTACTTAA 3840
 Oy 196 nAsp----- 197
 Db 3841 TGATGTAAATATCATATCTCTCTGCGGTCCTTGCGCTAGAACCTTCATATTACAG 3900
 Oy 197 ----- 197
 Db 3901 AAGAGATACATAGGCGCTGATGTTAGTTTGTACTCTCTCGCATCTCTTGGCA 3960
 Oy 198 -----GlyAsnArgGlyLe 202
 Db 3961 GGGTATTTGTACCAGAACTGATGTACAAATTAATGCGATGCTACAGGAACCGTGCTT 4020
 Oy 202 uGlyLysProGluAlaThrGluLeuAlaGlyLInPheGluMetLysSerAsnLysProPr 222
 Db 4021 AAGATATCCCGAGGACAGACAGACTGCTGGCAATTTGAGATGACGAGCAATTCACCC 4080
 Oy 222 oAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVa 242
 Db 4081 ACCCAATAGCCCACTCTCTCTGACCGCTGTCTAAAGTTATATTCACAGGGAAGCTGT 4140
 Oy 242 IValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259
 Db 4141 GGTCCCTGCTACTAGACAAAGAAAGTTATCTCTGAGCGATCAGAGGCTAGAAGGTTTGT 4200
 Oy 259 ----- 259
 Db 4201 TCATCATGACACCCGCTCATCATTAATTAACCAATTCCTGTTTACAATGTTCTTCCTATT 4260
 Oy 260 -----SerHisLeu 263
 Db 4261 ATGGATTAAGTGTATTAGTACTGTCATTAATTAACGAGAAATTTCTTCAGACGACCTACT 4320
 Oy 263 uLeuGlnLysArgLInPheThrHisSerHisArgVal----- 275
 Db 4321 TCTTGAGAAACCCCAATTCATCATCTTCACAGAGTCCAGTATCCATGCTTCACCC 4380
 Oy 275 ----- 275
 Db 4381 TACTCTTAAGCATTTTCTTAATGTCATGARGATATCTTAATCAAGCATACTTGCT 4440
 Oy 275 ----- 275
 Db 4441 TTGTTCTCATCAAAATTTGATTTGATGCTATGATGATCAAGCAAAATAGTTATGCTCA 4500
 Oy 276 -----Gln-Prometa 279
 III IIIIIII

Db 4501 TGTGTCTCCGTTTTTATTCGCACTAACCAAAAAATGATGTTCTGTGCAAGCAAGTGG 4560
 Oy 279 ILeuGlnGluValMetSerAspArgAspSerGluAspGluValAspAspAspValAla 299
 Db 4561 CGCTTGACAGTAATATGCTGTGACGGGATAGGAGATGAGTCACTGACATGCTTGCAG 4620
 Oy 299 sPpHeGluAspArg----- 303
 Db 4621 ATTTGAAATGCGCAGATTCATGATTTCTTCTGTGCTCATTAATATAGACAACAGA 4680
 Oy 304 -----GlnMetLeu 306
 Db 4681 AATGATATATGATGTAACCTCTAATAGCCTTTTGAACCTTAAGAAAGCTGCAGATGCTT 4740
 Oy 307 AspAspPheValAspValAsnLysAspGlyLysGlnPheMetHisLeuThrAsnSerPhe 326
 Db 4741 GATGACTTGTGTGATGTGATTAAGATGAAGAAAGCAATTCATCATCTTGTGAACTGTTT 4800
 Oy 327 ValArgLysGluArgVal----- 332
 Db 4801 GTTAGAAACAAAGGTA-ACACTTCTCTTACACATGAGACAGACAAAAAGACCTTANG 4859
 Oy 332 ----- 332
 Db 4860 TCTTACATTCATACCTCTGTAATGATTTGCTTATGGAACCTTGACGCTCAATTATGAT 4919
 Oy 333 -----IleAlaAspGlyHisLysSerThrAlaCysGluAlaPheSer 346
 Db 4920 TGTGATGTTTACAGGTTATACAGATGCTCATATTTCTTGGCAGATGACATTTTGA 4979
 Oy 347 ArgPheTyGlyLysGluLeuHisArgLysSerSerLeuPheThr----- 361
 Db 4980 AGATTTTACGAAAGAGTGTGACCGCTTATCATCTCTTGTGTAATATAGTACACC 5039
 Oy 361 ----- 361
 Db 5040 AATCATATACAGACACATACACTATCAATCTGTTTCTGTTCTGAAAAAATA 5099
 Oy 362 -----CysThrArgLeuPheLeuIleLysLeuThrAsnHisGlyLeuValAsp 377
 Db 5100 AAAATTTCCAGGTGTGGAGATGTTTGTGATTAATCTAATGAGAACCATGACTGTGCGAC 5159
 Oy 378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397
 Db 5160 TCAGCCACCATACAGACGCAATACCATCCCGAGAAATGGCGTATATGCTCAGACACC 5219
 Oy 398 ThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsn 417
 Db 5220 ACCACCACCAACAAACAAACAGTGTGATGCTCCAGTCAAGCTCAACCAACCAAT 5279
 Oy 418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsn 437
 Db 5280 AACATTTGTGRTCATCCCAATGACATTAACAAACAAAGAAATGTCACAAAGGACAT 5339
 Oy 438 AsnSerArgAspLysValIleLys 445
 Db 5340 AACAGAGAGACAAAGTATTA 5363
 RESULT 9
 US-09-890-220-6
 ; Sequence 6, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendell, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; FILE REFERENCE: Newburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; PRIOR FILING DATE: 1999-01-28

NUMBER OF SEQ ID NOS: 77
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO: 6
 LENGTH: 5895
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-890-220-6

Alignment Scores:
 Pred. No.: 1.79e-156 Length: 5895
 Score: 1659.00 Matches: 424
 Percent Similarity: 42.25% Conserved: 4
 Best Local Similarity: 41.86% Mismatches: 12
 Query Match: 69.76% Indels: 577
 DB: 33 Gaps: 14

US-09-890-220-2 (1-445) x US-09-890-220-6 (1-5895)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 1905 ATGTGTAGGAGAGATTTGCGCGGAAATCTCACCGAGAGAGATTTCAACGTATGAG 1964
 QY 21 AsnLeuLeuIleTyrCysIysProValArgLeuTyrAsnIlePheHisLeuArgSerIleu 40
 DB 1965 AATCTTGTGATATTTGTAACCTGTTCAGCTATATTAACATCTTTCACCTTCGCTCTCTA 2024
 QY 41 GlyAsn----- 42
 DB 2025 GGCAG-CGTATGATTTGGCCTTCCTCTCATCATTTTAGCTTAGTATATCTTTCATCTCC 2083
 QY 42 ----- 42
 DB 2084 TGTGTATATCACCACCTATATAGTTGAGTTGCTAAGCTGATTATGCTTCATCATGCC 2143
 QY 43 -----Pro 43
 DB 2144 GAGTGTGCTCTTTTGTCTCTAATGTTATTTAACACTGTGTGTGTGTGTCACGCCA 2203
 QY 44 SerPheLeuProArgCysLeuAsnTyrIleGlyAlaLysArgLysSer--- 62
 DB 2204 TCGTTCTGCGAAGAGCTTGAACTCAAAATTTGGGCAAGCGCAAAAGTA-TGC 2262
 QY 62 ----- 62
 DB 2263 GTTTCCTTGAATGAGTGGCAGAGATATGTTATTACTTCTAATATGAA 2322
 QY 62 ----- 62
 DB 2323 GCTGATGAACTATTTATCTTGTGTAGTATGACATATGATATGTTCTTCTTTC 2382
 QY 63 -----ArgSerThrGly 66
 DB 2383 TTCAATGCTATACCTATATTTTACAAATTTGTGTTTGTAGTCAAGATCTACAGGG 2442
 QY 67 MetValValPheAsnTyrIysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 DB 2443 ATGTAGTTTTCACATAAGAGTGTATATATATCAAAAGACTGAAGGTATGCT 2502
 QY 83 ----- 83
 DB 2503 TTTTCTGTTCTTCGACAAATTCGATGTCATAGTCTTCTCTAGATGTTGTTAT 2562
 QY 84 -----ValArgLysPheCysSerCysProPheCysSe 94
 DB 2563 TTACTATTTTTTCTGTATTTGTCACGAGATTAGGAGATTGTTCTGTCCATTTTCTC 2622
 QY 94 rMetLeuArgLysSerPheLys----- 101
 DB 2623 TATGCTATGTGGAGCTTCAGGTGGCAACTATTACAACTAGAGTTTCTCCGGGGCT 2682
 QY 101 ----- 101
 DB 2683 TTCAATCTAACACTGTGAATGCTACTGCGTTAATGCTATATACCTTCACTGTTTG 2742

QY 101 ----- 101
 DB 2743 TTACATATTTTGTGTTGTGTTGTTGCTTCTCTCTTTTAACTGCTGAGTGTGC 2802
 QY 102 -----G 102
 DB 2803 TTATCTGAGAAACATGTTCCAGTTCAGCTTCAATTCATGCTTGTGTCTATGAGG 2862
 QY 102 LysLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysPhe- 121
 DB 2863 GCTCCTCAATTCATTTGAAATTCATCATCATGATTTATTTGAAATTTGAGTCAAGATATC-T 2921
 QY 121 ----- 121
 DB 2922 GGTTTTATGCAATTTCTGTTTGTGCTATGCGCTGATGATGAGTTATGTTAAAAAG 2981
 QY 122 -----GluGluTyrGlnThrValAsnValSerValLysLeu 134
 DB 2982 GGTCTTCTCTATTTGAGCTTTTGGAGAAATACAGACAGTTATGTTCTGTAACCTTA 3041
 QY 134 snSerPheIlePheGlu----- 139
 DB 3042 ATTCCTCATATTTAGAGTCACTTAACTTGAATTTGGAAATCCCTATAGCTG 3101
 QY 140 -----GluGlnLysSerAspAsp 146
 DB 3102 GTGAAATTTGTTTATATTCATTCATCTTATTTGCTAGTAAAGAGATGATGAT 3161
 QY 147 LysPheGluProPheSerLeuCys----- 154
 DB 3162 AATTTGAGCCCTTCTCTCTCTG-CGTACTCTGAAACCCCTGATTAATACCTTAATA 3220
 QY 154 ----- 154
 DB 3221 GCAGTACTCTCTCTTCTGTCACTCTCTTAATTAATCAACCAACCAATGTTTGC 3280
 QY 155 ---SerLysProArgLysArgArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLys 173
 DB 3281 AGCTTGAAACCTCTGATGCGTAAAGAGAGGAGGAGAAATTAACCCAGGAGACTTAAA 3340
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGlnAsnGlyLysThr 193
 DB 3341 GATGCTTTTACCGTTGATTCACCCAGTTTGTAAATGCGCAGAAATGGAATTCGC 3400
 QY 194 LeuLeuAsnAsp----- 197
 DB 3401 CTGCTGAAATGATGTAATAATCAATCTTCTGTGTGATTCGTTGCTGATAGAACTTC 3460
 QY 197 ----- 197
 DB 3461 ATTTTACAGAAAGATACAAATGCTGATTTGTTTGTGTTTCTCTCTGCAATTC 3520
 QY 198 -----GlyAs 199
 DB 3521 TTTCTGTAGGCTATTTTACAGAACTGATGTACAAATTTATGCACTGCTACAGAAA 3580
 QY 199 nArgGlyLeuGlyLysProGluLysArgLysLeuLysGlnPheGluMetThrSerAs 219
 DB 3581 CGGTGGTTTATGAGATATCCAGAGCAAGAGCTTCTGACACATTTTGATGATCTACAGCA 3640
 QY 219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 DB 3641 CATTCACACACCTTACCCCACTCTCTGACGCTGTGCTGTAAGTTATATTAACAAAC 3700
 QY 239 rGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
 DB 3701 CGAAGCTGTGGTCCCTCTCTACTAAGACAAAGAAATTTATCTGCTGAGCATCAGAGGCTAG 3760
 QY 259 g----- 259
 DB 3761 AAGGTTTGTTCATCATGACACCCGTCATCATATTAATACATACCTGTTTACAAATGTT 3820

```

OY 260 -----Se 260
DB 3821 CTTCATTATGATTAAGTTTACTGCTACGATTTAACCAGAAATTTCTTCAG 3880
OY 260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
DB 3881 CCACCTACTCTTCAGAAACGCCAATTTATCATCTCTCACAGATCCAGGTGATCCAGT 3940
OY 276 ----- 276
DB 3941 TCCCTCACCCTCTTCTTAGGCATTTTCTTAAATGCTCATGATGATATCTTATCAAGC 4000
OY 276 ----- 276
DB 4001 ATACTGTGTGTCTCATCTAAATTTGTATTTTGTATCTGTATGTATCAAGCAAAAA 4060
OY 276 ----- 276
DB 4061 ATTATGTCATGTGTCTCCGTTTATGTCACATAACCAAAAACTCATGTTTCTGTGA 4120
OY 277 -----PrometAlaLeuGlnGlnValMetSerAspArgSerGlnAspGluValAspAs 295
DB 4121 CAAGCCAAATGGCGCTTGACCAAGTAATGCTGATCGGATAGCGAGATGAAGTCATGA 4180
OY 295 PASPYAlaAlaAspPheGluAspArg----- 303
DB 4181 CGATGTTCAGATTTTGAAGATCGCAGATTTCCATGATTTCTTCTGCGTTCAATAG 4240
OY 303 ----- 303
DB 4241 TAGGCACAGAAATGGTATACGATGTAACCTGCTTAAGCTTTGAAACTTAAAAAAGC 4300
OY 304 ---GlnMetLeuAspAspPheValAspValAspLysAspGlnPheMetHisLeuT 323
DB 4301 TCGAGTGTCTGATGCTTGTGGATGTAAATGAATGAAGCAATTCATGCTCTT 4360
OY 323 rPAsnSerPheValArgLysGln----- 330
DB 4361 GGAAGCTCGTTTGAAGAAAAAAGTAAGTACTTCTCTTACACTTGAACACACACAAAA 4420
OY 330 ----- 330
DB 4421 AGACCTTATGCTTACATTCATACCTGTCTAATGATCTGTATGGAAGCTTGAAGCT 4480
OY 331 -----ArgValIleAlaAspGlnHisIleSerThrAlaCysGln 343
DB 4481 CAATATGATGATGATGTTGCAAGGTTAAGCAGATGTCATATCTCTTGGCATGTGA 4540
OY 344 AlaPheSerArgPheTyrGlnLysGlnLeuHisArgTyrSerSerLeuPheTyrP 361
DB 4541 GATATTTCAAGATTTTACGAGAAAGGTTGCACTGTACTCATCACTCTTCTGTAATAT 4600
OY 361 ----- 361
DB 4601 AAGTACCCAAACATATACAGACACATACATCAATTTTCTGTTTTCTGA 4660
OY 362 -----CysThrArgLeuPheLeuIleLysLeuThrPAsnHisGln 374
DB 4661 AAGAAAAATTAATAATTTCCAGGCTGTGGAGATTTGTTGATTAACCTATGGAACATG 4720
OY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlnAsnCysArgAsn 394
DB 4721 ACTGTGCACTCAGCCACATCAACAACATGCAATTCCTGCAATATGCGGTATATAC 4780
OY 394 rSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsn 414
DB 4781 CTCAC-----GTCACCTACACAAACAACAGTGTGATCATCCAGTACCAACAAAC 4834
OY 414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnValAspAs 434
DB 4835 CAACACAAATTAACATTTGTGATCATCCGATGACATTAATAAACAACAAGATGTTGACA 4894
OY 434 nLysAspAsnAsnSerArgAspLys 442

```

```

DB 4895 CAAGGACAAATTAACAGACAGACAG 4919
RESULT 10
US-09-534-859-587
; Sequence 587, Application US/09534859
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Last, Robert L.
; APPLICANT: Levin, Irene M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
; FILE REFERENCE: 38-10(15493)B
; CURRENT APPLICATION NUMBER: US/09/534, 859
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 1127
; SEQ ID NO 587
; LENGTH: 201471
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-534-859-587
Alignment Scores:
Pld. No.: 2,39e-154 Length: 201471
Score: 1659.00 Matches: 424
Percent Similarity: 42.25% Conservative: 4
Best Local Similarity: 41.86% Mismatches: 12
Query Match: 69,768 Indels: 577
DB: Gaps: 14
US-09-890-220-2 (1-445) x US-09-534-859-587 (1-201471)
OY 1 MetCysArgGlnAsnAsnCysArgAlaLysSerSerProGlnGlnValIleSerThrAspGln 20
DB 39570 ATGTGAGGACAGAAATGTCGCGGAAATCTCCACCGAGGAAGTATTTCAACTGATGAG 39629
OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 39630 AATCTCTGATATATGTAACCTGTCGACTATATATACATCTTCCACTGCTCTCTCA 39689
OY 41 GlyAsn----- 42
DB 39690 GGCAG-CGATATGATTTGGCTTCTCTCATCATCTTTAGCTTAAATCTTTCATCTCC 39748
OY 42 ----- 42
DB 39749 TGTGTAGATCACCCACTAATATGATTTGATTTGCTAGCTGATATGCTGACTCATGGC 39808
OY 43 -----Pro 43
DB 39809 GAGTGTGCTCTCTTTTGTCTCTAATGTAATTTGAATCTGTTGTTGTTGTCAGCCA 39868
OY 44 SerPheLeuProArgCysLeuAsnTyrLysIleGlnAlaLysArgLysSer----- 62
DB 39869 TCGTTTTCGCAAGATGCTTGAACACAAATTTGGGCAAGGCAAAAGAAAGTA-TGC 39927
OY 62 ----- 62
DB 39928 GTTTCCTTGAATGATGTTGCCACAGTGATATGTAATTTATCTTACTTCTAATATGAA 39987
OY 62 ----- 62
DB 39988 GCTGATGAACATTTATCTTGTGTGATGATATGACATATGAATGATGTTCTTCTTGG 40047
OY 63 -----ArgSerThrGly 66
DB 40048 TTCATGCTATACACTTATATTTTACAAAATTTGCTTTTGTCTTATAGTCAAGATCTACTGG 40107
OY 67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83

```

Db 40108 ATGGTACTTTCAACTATAGATTTGTAATATACATTCAAAAGACAGAGGTTAGTCT 40167
 QY 83 ----- 83
 Db 40168 TTTTCTGTCTTGCACAAAATTCGATGCAATGCTATGTTCTCTAGATGATTTGTTAT 40227
 QY 84 ----- -ValargluaspCysSerCysProphCysSe 94
 Db 40228 TTACTATTTTTTCTGTATTGTGCAGCAGTTAGGGAGATTTCTTGTCCATTTTGTCTC 40287
 QY 94 melleucysglySerPheLys----- 101
 Db 40288 TATGCTATGTGTAGCTTCAAGGTGGCAACTATTACAACTGAGGTTTCTCCGGGGCCT 40347
 QY 101 ----- 101
 Db 40348 TTCAATATCAACACTGGAATGCTACTGCCGTTTAAAGCTATATACATTTCAGTGTGG 40407
 QY 101 ----- 101
 Db 40408 TTACATATTTTTGTGTGTGTGTCTTCTTCTTTTAACTGCTGAGTGTGC 40467
 QY 102 ----- -G 102
 Db 40468 TTATCTGAGAAACATGTCCTGAGCTTACATTCATTCCTGTGTGTATGACAG 40527
 QY 102 ----- 102
 Db 40528 GGTGCAATTTTCATTTGAATTCATTCATGATTTATTTGAATTTGATTCAGTCAAGCTATG-T 40586
 QY 121 ----- 121
 Db 40587 CGTTTATGGAATTTCTGTGTTGCTATGCCCTTACTAGTAAGCTTATAGTTAAAAAG 40646
 QY 122 ----- GluGluTyrGlnThrValAsnValSerValLysLeu 134
 Db 40647 GGTCTTCTCTATTTGCTAGCTTTTGGAAATACACAGACAGTATATGTTCTGTAAACCTTA 40706
 QY 134 snSerPheLlePheGlu----- 139
 Db 40707 ATTCCCTCATATTGAGGTCAGTACTTAAACTGTTAATTTGGAAATCCTTATAGCTG 40766
 QY 140 ----- GluGluGlySerAspAsp 146
 Db 40767 GTGAAAATTTGGTTTATTTCCATCTTATTTGTACTAGGAAGAAAGATGATGATGAT 40826
 QY 147 LysPheGluProPheSerLeucys----- 154
 Db 40827 AAATTTGAGCCCTTCTCTCTG-GTAACTCTCAGAACCCCTTGATTTAAATACCTTAATA 40885
 QY 154 ----- 154
 Db 40886 GCAATACCTCTGCTTTTCTGTCAGTACTCTCTATTAATCAACACCAATGTTTGC 40945
 QY 155 ---SerLysProAlaTyrGlnAspArgGlnArgGlyArgAsnAsnThrArgLeuLys 173
 Db 40946 AGCTCGAAACCTCGTAAGCTAGACAAAGAGGGGAGAAATTAACACGAGGACTTAA 41005
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGlnAsnGlyLleThr 193
 Db 41006 GTATGCTTTTACCGTTGGATTCACCCACTTTAGCTAATGACAGAAAATGGAATTC 41065
 QY 194 LeuLeuAsnAsp----- 197
 Db 41066 CTGCTGATGATGTGTAATCAATCTTCTCTGTGTGTATTCGTTGTGCTTGAACCTTC 41125
 QY 197 ----- 197
 Db 41126 ATTTTACAGAAGAATACAAATGCTCGATTGTTAGTTTGTACTCTCTCCGCAATTC 41185
 QY 198 ----- -GlyAs 199
 Db 41186 TTCTGTGAGGTAATGTTACAGAACATGATGTACAAAATTAATGCACTGCTACAGGAAA 41245

QY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
 Db 41246 CCGTGTTTATGATATCCCGAGGCAACAGACCTTGGCGCAATTTGAGATGATGAC 41305
 QY 219 nIleProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 Db 41306 CATTCACAGAGCCATAGGCCACTCTTCTGTGAGCGTGGTCAAGATTATTTAAACAAC 41365
 QY 239 rGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluAlaSerGluAlaAr 259
 Db 41366 CGAAGCTGTGTCCCTGCTACTAAGACAAAGAAATTTATCTGTGACGATCAGAGGCTAG 41425
 QY 259 g----- 259
 Db 41426 AAGTTTGTTCATCATGACACCCGTCATCATATATACAACTGTGTACAAATGTT 41485
 QY 260 ----- -se 260
 Db 41486 CTTCCTATTATGATATAGTGTATTACTGTAAGCCATATTAACCGAGAAAATTTCTTCAG 41545
 QY 260 rHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
 Db 41546 CCACACTCTTCTGAGAAAGCCAAATCTATTCATTCACAGAGTCCAGGTGATCAAGT 41605
 QY 276 ----- 276
 Db 41606 TCCCTCACCTACTTCTTAGGCATTTTCTTAAATGCTCATGATATCTTATCAAGC 41665
 QY 276 ----- -se 276
 Db 41666 ATACTGTGTTGTTCTCATTAATTTGATTTGATTCGTATGATCAACGCAAAAA 41725
 QY 276 ----- 276
 Db 41726 ATTAATGCCATGTGTCTCCGTTTATTTGCACTAACCAAAACTGCATTTCTGTGCA 41785
 QY 277 ---PrometaLeuGlnGluValMetSerAspArgAspSerGluAspGluValAspAs 295
 Db 41786 CAAGCCAATGGCGCTTGAGCAATTAATGTCTGATCGGAGATGCGAAGATGAAGTGA 41845
 QY 295 pasPValAlaAspPheGluAspArg----- 303
 Db 41846 CGATGTGCAGATTTTGAAGATCCGACGATTCATGATTTCTTCTGCTTCATTAAG 41905
 QY 303 ----- 303
 Db 41906 TAGCCAAAGAAAATGATATGATGATGATGCTGTAATGGCTTTGAAACTTAAAAAGC 41965
 QY 304 ---GlmMetLeuAspAspPheValAspValAsnLysAspGlyLysGlnPheMetHisLeu 323
 Db 41966 TGCAGATGCTTGATGACTTTGTGATGTGAATTAAGTAAAGCAATTCATTCATCTTT 42025
 QY 323 rPAsnSerPheValArgLysGln----- 330
 Db 42026 GGAATGCTTTGTAAAGAAAACAAGGTAAGTACTGCTCTTACACTTGAACACACAAAA 42085
 QY 330 ----- 330
 Db 42086 AGACTTATGTCTTACATTCATACCTGCTCTAAATGATTTCTGTATGAACTTTGAGCT 42145
 QY 331 ----- ArgValIleAlaAspGlyHisLysSerThrPalacyGlu 343
 Db 42146 CAATATTATGATGATGTTTGCAGGGTATAGCAAGATGTCATATCTTGGGCAATGTAA 42205
 QY 344 AlaPheSerArgPheTyrGlyLysGluLeuHisArgLysSerSerLeuPheThr----- 361
 Db 42206 GTATTTTCMAATTTTACGAGAAAGAGCTTCACCTGTTACTCATCACTTTCGTGAATAT 42265
 QY 361 ----- 361
 Db 42266 AAGTACCAACAATATACAGACACATTAACATATCAATTTGTTGTTTCTGTA 42325

Query Match: 49.03% Indels: 46
DB: 42 Gaps: 9

US-09-890-220-2 (1-445) x US-10-219-999-5373 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 164 ATGTCCCGCAAAATTTTCCTCCGTACACCATCCGCGGGAAGAAATATTCACCTGCTGATG 223

QY 21 AsnLeuLeuIleTyrCysLysProValAlaGluLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 224 AGCCCTTGATTTATTTAGCAAGCCCTGTGAACTTTACAAATATTCCTACCGCGCTGCTTT 283

QY 41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 284 CAAATTCCTCTTTCTTCTTATGAGATCTTGCCTTAAATAAGCAACCCCTAAAG 343

QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 344 AGCTTGAGAGCA---GGAATGTGATTTCAATTAAGGACCGCTACCAACATTCCTGCA 400

QY 81 LysThrGluValArgLysAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 401 AGACTGAAGTACCGAAGACCTTTCTTCTCCGTTTGTGATGCAAGTGTGAAACTTT 460

QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 461 AAGGCTTTCGATTCATCTTGTTCATCATCATGATCATTCACCTTGAGTTCTGCTG 520

QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 521 ACTGAAGATACCAAGCAAGTGAATGCTCTGTAAATAATGAT---ATAATGAGATCA 574

QY 141 GluGlySerAspAsp-----AspLysPheGluProPheSerLeuGlySerLysPro 157
DB 575 GAGATGTTGCTGATGAGTAATTCACATGCGCAACCTTCTCTTCTGCTCAACACT 634

QY 158 ArgLysArgArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
DB 635 CGAAACAGTAAAGAAAGACGCTGTCAAATGAAAGCGCAACATGTAATTCCTG 694

QY 178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAsp 197
DB 695 GAGTGGACTACCA-----GAAAGCATACATTAATGATTTCTACAAAAAGATGAT 745

QY 197 ----- 197

DB 746 GATATCTTATCTGCAAGAGAGAAATGTGCTAGAACATCTCTGATGAGAGATTTTC 805

QY 198 -----GlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu 210
DB 806 CCTAGTGAAGAAATGACGAGAGAAATTTGTCCTGATCATCTGCGCACCATGAGCAAC 865

QY 211 AlaGlyGlnPheGluMetThrSerAsnIleProPro-----AlaIleAlaHisSer 227
DB 866 CTGAGACATGAGTGAATCCAGTTTCCAGATTCAGAGTTCATTCATCCATGCCCAATCT 925

QY 228 SerLeuAspAlaGly---AlaLysValIleLeuThrSerGluAlaValValProAlaThr 246
DB 926 TCTGTGACCTCGAATGTAGTAATCAATGTAAAGATGTAATCTCTGCTGCT--- 982

QY 247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLys 266
DB 983 AAAACAAAGACCTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1042

QY 267 ArgGlnPheTyrHisSerHisArgValGlnPrometAlaLeuGluGlnValMetSerAsp 286
DB 1043 AGACTTTCCTTCACTCACACAGAGTCCAGCTATGAGCATGAGCAACAGAGTATATGAG 1102

QY 287 ArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeu 306
DB 1103 CGGTATGATGAAGACAAAGTGAATGACGATTCGAGATCTTGAAAGATGAAGAGCTT 1162

QY 307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPhe 326
DB 1163 GACGATTTTGTGATGATTTTCCAAAGATGAAAAACACCTCATCTCTGGAACCTCTTT 1222

QY 327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSer 346
DB 1223 ATGAGGACCAAAAGGCTGCTAGAGATGATGATGATGATGATGATGATGATGATGAT 1282

QY 347 ArgPheTyrGluLysLeuLeuHisArgTyrSerSerLeuPheThrCysThrArgLeuPhe 366
DB 1283 AGCTTCATGAGAAAGAGCTGATCTCATCTCCAGCTTATTTTGGTGGAGGTATTC 1342

QY 367 LeuIleLysLeuThrPheHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
DB 1343 ATGATCAAACTTGGATATCATGCTGCTTGTGATGCTGTACATGACAACTGTACATA 1402

QY 387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerVal 406
DB 1403 GTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1447

QY 407 AspArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisPro 423
DB 1448 TGAAGACATTAAGAAAGTACTCAACAAACAAATATACAGTCTTCTCCG 1498

RESULT 13
US-09-339-947A-2
; Sequence 2, Application US/09339947A
; GENERAL INFORMATION:
; APPLICANT: YOSHIDA, Nobumasa
; APPLICANT: YAMAI, Yukihito
; APPLICANT: KATO, Yoshihiro
; APPLICANT: HIRATSUKA, Junzo
; APPLICANT: TAKAHASHI, Shigeru
; APPLICANT: MIWA, Tatsushi
; TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
; FILE REFERENCE: 032735-001
; CURRENT APPLICATION NUMBER: US/09/339,947A
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: JP 1998-180065
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: JP 1999-179043
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (310)..(2142)
; NAME/KEY: misc_feature
; LOCATION: (1650)..(1655)
; OTHER INFORMATION: BamHI recognition site
; NAME/KEY: misc_feature
; LOCATION: (1984)..(1989)
; OTHER INFORMATION: SphI recognition site
US-09-339-947A-2

Alignment Scores:
Pred. No.: 2,23e-82 Length: 2280
Score: 923.00 Matches: 216
Percent Similarity: 43.23% Conservative: 52
Best Local Similarity: 34.84% Mismatches: 108
Query Match: 38.81% Indels: 244
DB: 17 Gaps: 9

US-09-890-220-2 (1-445) x US-09-339-947A-2 (1-2280)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 310 ATGTCCCGCAAAAGATCCGCTGCTATTTCTGAGAGAGAGAGATTCCTGCTGAAAGAG 369


```

OY 21 AsnLeuLeuIleTyrcyslysprovalArgLeuTyAsnIlePheHisLeuArgSerLeu 40
   :|:|:|
Db 370 ACCTTGCTGCTATTGTCAGACCTGTCGAACTGTCACATATTCATTCACGCCGCTCTAT 429
OY 41 GlyAsnProSerPheLeuProArgcysLeuAsnTyrlsIleGlyAlaLysArgLysArg 60
   :|:|:|
Db 430 AGAATCCCTGTTCTTCAGCCAGATGTTGATTAAGATGAGCAAAACATAAAG 489
OY 60 -----
Db 490 AGAATACAAATGACTGATTCCTCTCGGGCGCTATAGATCTGGGGTACAAATCMAAA 549
OY 60 -----
Db 550 TTATTCCTCTGTATATTTTGGTGGCAAGACTGTTCTCTAAGCCTGCTGATAT 609
OY 60 -----
Db 610 TCTGCAATATAGATTGCTGAGCATGATCTTAAGTGTGATGGGGGTGATGA 669
OY 60 -----
Db 670 GTTAGTCAAGCCCAAGCAACTTCTTCTGATGATGATAGATGAGCTGCAATGAGGCA 729
OY 60 -----
Db 730 AAATCAGGATCCTGCTATCTTGTTCATCAGCTTCTGTCGCGCAAAATTCATTT 789
OY 60 -----
Db 790 GGCATGATTCAGGCAAGATTTCATGAGAAATATAGAGACATTTGTTATGAGCAAA 849
OY 60 -----
Db 850 ATACCTGCAATCAGTATGCTGCTGCGCAAAATCACCMAAATGAGACTGGGAGAG 909
OY 61 -----
Db 910 AGATGACACACAGCTCTCTGTTGTAAGACAGCCCTGTTATAGCTAAGTAAAGTCAAG 969
OY 64 Ser -----
Db 970 AGTCAAGGAAAAGTGTCTGATTCAGTGCACCAATCATCACTGAGCTCTCG 1029
OY 64 -----
Db 1030 CAGCAAGTGCAGATCAGCATATCTGAGAGAAGATTGGTCAACGAAAAATCTCTTAT 1089
OY 64 -----
Db 1090 AGTTCATTTTCATATAATGACATCTCTCTCTCTCTCTCTCTGCAAAATATAGGTTGAGA 1149
OY 65 ThrGlyMetValValPheAsnTyrlsAspCysAsnAsnThrLeuGluLysThrGluVal 84
   :|:|:|
Db 1150 ACAGAAATATGATGTTTCAACATCTGATCTATTAACAAATTCAGAGAGCTGAAAGTA 1209
OY 85 ArgGluAspCysSerSerPheCysSerMetLeuGlySerPheLysGlyLeuGlu 104
   :|:|:|
Db 1210 ACTGAAGACTTTCTTCTGATCTCTGCTTAAGAAATGAGGCAAGTTTCAAGGCTGAGA 1269
OY 105 PheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluLys 124
   :|:|:|
Db 1270 TATCATCTGCAATCAACCCAGATCTCTCAATTCGATTTGGGTGATGAGAAATTT 1329
OY 125 GluThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluGlySerAsp 144
   :|:|:|
Db 1330 CAGGCGGTAAATGTTCTCCCAAGACTGAGACAAATGATTCAGAGTTAAATGAGATGAC 1389
OY 145 ---AspAspLysPheGluProPheSerLeuGlySerLysPro---ArgLysArgArgGln 162
   :|:|:|
Db 1390 GTTGACCCCAAGCAGCAACTTCTTTTCTTCTCAAAAAAATTCAGAGGAGAGCAAA 1449

```

```

OY 163 ArgGlyGlyArgAsnAsnThrArgArgLeuLysValcysPheLeuProLeuAspSerPro 182
   :|:|:|
Db 1450 AAGAGTCAGTACGAGGATCCAGGCA-----ThrgluasnGlyIleThrLeuAsn 1482
OY 183 SerLeuThrAsnGly-----ThrgluasnGlyIleThrLeuAsn 196
   :|:|:|
Db 1483 CATCTTGATTTGGTGGAGGTGCTAGATTAAGCTATGATGCTATCTGTTAGAAAGT 1542
OY 197 AspGlyAsnArg-----GlyLeuGlyTyProGluIleThrGluLeuIleGlyGln 213
   :|:|:|
Db 1543 GAGAAAGCGCAATACCACTGGAAGCATTC-----GAAATATGGGGCT 1590
OY 214 PheGluMetThrSerAsnIlePro-----ProAlaIleAlaHisSerSerLeu 229
   :|:|:|
Db 1591 GCTGAGTCTGTCAAAAGATTCTCTGCGACAGATCTCTGACAGCGTCAATCATGTGG 1650
OY 230 AspAlaGlyAlaLysValIleLeuThrSerGluIleValAlaValProAlaThrLysThrArg 249
   :|:|:|
Db 1651 GATCCAGATTTATGTCAGTGCATGCTGGAAGTACATGTTCAGATTGCAAAAGCAGG 1710
OY 250 LysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArgLysPhe 269
   :|:|:|
Db 1711 AAAATATCTATGAACCGTCCGACTGAGAACCGAAGCCTCTTCGAAAGACAGCTTC 1770
OY 270 TyrHisSerHisArgValGlnPrometAlaLeuGluGlnValMetSerAspArgSer 289
   :|:|:|
Db 1771 TTCACGTCTATCGAGCTGACGATGGCTGTAGAACAAAGTCTTGGACCGGATTA 1830
OY 290 GluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPhe 309
   :|:|:|
Db 1831 GAGAGTGAATGATGATGATGTGCGAATTTGAAAGAGTGTGATGATTC 1890
OY 310 ValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPheValArgLys 329
   :|:|:|
Db 1891 GTTGATGTGACTAAAGATGAGAAACAGATGTGACATGTGGAACGCTTGTGAGGAAG 1950
OY 330 GlnArgValIleAlaAspGlyHisIleSerThrAlaCysGlnLalaPheSerArgPheTy 349
   :|:|:|
Db 1951 CAGCGATATTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2010
OY 350 GluLysGluLeuHisArgTySerSerLeuPheThrPysThrArgLeuPheLeuLys 369
   :|:|:|
Db 2011 GGACCCATCATGCTGTCGAACCCGACCTGATTTGGTGTGAGAGTGTGTTATGGGAA 2070
OY 370 LeuThrAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleGluGln 389
   :|:|:|
Db 2071 CTGTGGAACCAAGGCTTCTTGTGATGCCGAACATGACAACTGTAATACCTTCTCGAA 2130

RESULT 14
US-60-303-460-9
: Sequence 9, Application US/60303460
: GENERAL INFORMATION:
: APPLICANT: Allen, Steve
: APPLICANT: Butler, Karlene
: APPLICANT: Niu, Xiping
: TITLE OF INVENTION: Flowering Regulating Polynucleotides and Polypeptides
: FILE REFERENCE: BRI493 US PVY
: CURRENT APPLICATION NUMBER: US/60/303,460
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 9
: LENGTH: 2346
: TYPE: DNA
: ORGANISM: Helianthus sp.
US-60-303-460-9

Alignment Scores:
Pred. No.: 8.68e-81
Score: 907.50
Percent Similarity: 43.19%
Best Local Similarity: 33.38%
Query Match: 38.16%

Length: 2346
Matches: 218
Conservative: 64
Mismatch: 120
Indels: 251

```


TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
 FILE REFERENCE: 032735-001
 CURRENT APPLICATION NUMBER: US/09/339,947A
 PRIOR FILING DATE: 1999-06-25
 PRIOR APPLICATION NUMBER: JP 1998-180065
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: JP 1999-179043
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 9
 LENGTH: 2248
 TYPE: DNA
 ORGANISM: *Oryza sativa*
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (86)..(1897)
 NAME/KEY: misc_feature
 LOCATION: (36)..(41)
 OTHER INFORMATION: NheI recognition site
 US-09-339-947A-9

Alignment Scores:
 Pred. No.: 1 03e-80 Length: 2248
 Score: 906.50 Matches: 209
 Percent Similarity: 42.688 Conservative: 59
 Best Local Similarity: 33.288 Mismatches: 109
 Query Match: 38.128 Indels: 251
 Gaps: 7

US-09-890-220-2 (1-445) x US-09-339-947A-9 (1-2248)

QY 1 MetCysArgGlnAsnGlySerProGluGluValIleSerThrAspGlu 20
 DB 86 ATGGCCGACCAAGGCGCTGCTCTCCGATGACGACCTTGCAAGAA 145
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 146 AGCTGCATTAATGCAAGCGGCGGTGATTAATATCATTCACGGCCGATCCATT 205
 QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 206 AAAAATCCCTCTTCTTCTCAAGATGCCCTTTTCAAGATTCAAGCAAGCAAGCAAG 265
 QY 61 LysSerArgSerThr----- 65
 DB 266 AGAGCCCTGATTAACCATATCACTTCTGTGGAGCACAATAAAGAACTCGGGCACAATA 325
 QY 65 ----- 65
 DB 326 ATCTTCCTCTTATGTTCTGTAGTAGACTACTAATAATGTTTCACTTGAAGGCAT 385
 QY 65 ----- 65
 DB 386 TCTCGATATATGATTGATTCAGTGTGCTGTTGTTGACTCTTTTCAATGATTTGGAAT 445
 QY 65 ----- 65
 DB 446 AAAGACTACACTGAACCAATTCCTGATCTCTGATGTGAAGAACTTAGCACTCCCGA 505
 QY 65 ----- 65
 DB 506 GCTTGCAGCCTTAATATTATCTTATCAGCTGTGGAGCAGCTGAGCAAACTTTTGATGAC 565
 QY 65 ----- 65
 DB 566 AATTAAGTCTTGGGAACCATGTGAAGGCTCTACTCTCAAAAGCTTGAAGGAAGTGT 625
 QY 65 ----- 65
 DB 626 TTCTGGGTAATAATACCAATCGATCTTCTGCTTCATCTTGGGAATTTGTGTGAGCTTA 685

QY 65 ----- 65
 DB 686 AGTTGGAGCATACCGTGGAAATGCTTCCACGGTGTGATGACCAAGCAAGCTTCTAGAG 745
 QY 65 ----- 65
 DB 746 CCAAAATTTCTGAGAGATGACAGTGTGCTTACATTTTGTCTGAGAAGGTGATGCTACT 805
 QY 65 ----- 65
 DB 806 GGTTCATTCAACGTCAAGTACATATCTGTCTCAAGAGCGTGTGCAAAAGCATGTTC 865
 QY 65 ----- 65
 DB 866 GAGTCTCCTTATAGTGTATTATATATGATGATGACACCTTCGTCATTGACATATT 925
 QY 66 ----- 925
 DB 926 ATTAAGCTTGAATCTGCAATGCTCTTTTACTACAAATACATTAATATCTGCA 985
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 DB 986 AAACCCGAGTCACTGAAAGATTTTCTGCGCATTTTGTGTGATGATGACAGCTTT 1045
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLeu 120
 DB 1046 AAGGCTTGAAGATGATCACTTAACGATCGCATGACCTTTCCATTTATGATTTGGATA 1105
 QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu 139
 DB 1106 TCTGAAGAGTCCAGGCTGTTAATGTTAATGCTAAGACATGATCTTGGAGAAACAGCTT 1165
 QY 140 ---GluGluLysSerAspAspLysPheGluProPheSerLeuGlySerProArg 158
 DB 1166 TTGGCTGAGGAGAT---GATCCAAAGCATCAAAATTTTCTGATCCGCAAGATTTAG 1222
 QY 159 LysArg---ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
 DB 1223 AACCGTAAAGAGGAGAAATCTCAAGTGAATAATGAGCATGATCATCATATTGTG 1282
 QY 178 ProLeuAspSerPro---SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsn 196
 DB 1283 GATTGAGGATCACTGGAAGATGCCAGCGAGATCTGAAGACATTAAGTCAAGAGGAA 1342
 QY 197 AspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMet 216
 DB 1343 AATGTAGT----- 235
 QY 217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal--- 235
 DB 1352 ----- 235
 QY 236 -----IleLeuThrSerGluAlaValAlaProAlaThrIleLysThrArgLysLeuSerAla 253
 DB 1394 GGTAGCAATCTTTTACGACCAACAGTGTTCACGTTTGGGAGCAAGAAAGTGTCTGTT 1453
 QY 254 GluArgSerGluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHis 273
 DB 1454 GAAAGGCTGATCCCAAAATTCGCAAGCTCTCAAAAGCAAGCTTCTTATTTCTCAC 1513
 QY 274 ArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVal 293
 DB 1514 AGGCTCAAAACCAAGGCTGAGCAAAATTTTCTCAGATCGATGATGAGATCAATTT 1573
 QY 294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
 DB 1574 GATGATGATCATGCTGATTTTGAAGATAGAAAGATCTTGATTTGTTGATGTACA 1633
 QY 314 LysAspGluLysGlnPheMetHisLeuTyrPheSerPheValArgLysGlnArgValIle 333
 DB 1634 AAAGCGAGAAACTTATATGATATGTGAATTTCTTGTTCGGAACCAAGGCTACTA 1693
 QY 334 AlaAspGlyHisIleSerThrPheLysGluAlaPheSerArgPheTyrGluLysGluLeu 353

```

Db      1694 GCGGATGGCATATTCCTGGGCATCGAGCATTCGCACTTCATGACAGAACTT 1753
OY      354  HIsArgTYrSerSerLeuPheTrpCysTrpPArgLeuPheLeuIleLysLeuTrpAsnHis 373
Db      1754 GTACAAATCCAGCTCTACTATGAGGTGGAGGTTTTTATGTCAAACTCTGGAACCCAC 1813
OY      374  GlyLeuValaspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393
Db      1814 AGTCTACTGGATGGCGGCATGATGCTGCACACATCTTGAAAGGCTACCTGAAC 1873
OY      394  SerSerAspThrThrThrAsn 401
Db      1874 GGAAGCTCGATCCAAAGAAAAAT 1897

```

Search completed: June 20, 2003, 02:18:49
 Job time : 3231 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:51:48 / Search time 956 seconds
(without alignments)
3236.070 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRONCRASPEEIVSYDE.....INNNKNNVDKNNSHDKYIK 445

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7816032 seqs, 3476047932 residues

Total number of hits satisfying chosen parameters: 15632064

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_SPOOL/US09880220/runat_13062003_144853_16556/app.query.fasta.1.583
-DB=Pending_Patents_NA_New -OFMT=fastap -SUFFIX=trnp -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09880220.ecgn.1.1.844.etrnat.13062003.144853.16556 -NCPU=6 -ICPU=3
-NO_MAP -LARGEREQUY -NEG.SCORES=0 -WAIT -DSPRLOCK=100 -LONGLOCK
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCIT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*
13: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1166	49.0	1722	US-10-425-114-11348	Sequence 11348, A
2	1166	49.0	1838	US-10-424-599-111422	Sequence 111422, A
3	833	35.0	1841	US-10-425-114-31883	Sequence 31883, A
4	653	27.5	2640	US-10-231-778-6	Sequence 6, Appl1
5	648.5	27.3	2700	US-10-437-963-91042	Sequence 91042, A

Result No.	Score	Query Match	Length	ID	Description
6	617.5	26.0	1004	US-10-437-963-11609	Sequence 11609, A
7	575.5	24.2	1070	US-10-424-599-141143	Sequence 141143, A
8	492	20.7	2439	US-10-425-114-30709	Sequence 30709, A
9	479	20.1	716	US-10-424-599-111421	Sequence 111421, A
10	472.5	19.9	443	US-09-837-604B-13507	Sequence 13507, A
11	472.5	19.9	443	US-09-837-604A-13507	Sequence 13507, A
12	435	18.3	558	US-10-389-048-21629	Sequence 21629, A
13	348.5	14.7	554	US-09-531-113-55697	Sequence 55697, A
14	348.5	14.7	554	US-09-531-113-55697	Sequence 55697, A
15	292	12.3	531	US-09-837-604B-940	Sequence 940, App
16	292	12.3	531	US-09-837-604A-940	Sequence 940, App
17	271	11.4	322	US-10-424-599-40645	Sequence 40645, A
18	264.5	11.1	6458	US-10-231-778-7	Sequence 7, Appl1
19	256.5	10.8	433	US-10-424-599-73456	Sequence 73456, A
20	255.5	10.7	453	US-09-837-604B-6414	Sequence 6414, App
21	255.5	10.7	453	US-09-837-604A-6414	Sequence 6414, App
22	253.5	10.7	453	US-10-437-963-97425	Sequence 97425, A
23	221	9.3	3235	US-10-235-25585	Sequence 25585, A
24	204.5	8.6	3900	US-10-144-771-9790	Sequence 9790, App
25	191.5	8.1	250	US-09-531-113-40606	Sequence 40606, A
26	191.5	8.1	250	US-09-531-113-40606	Sequence 40606, A
27	190	8.0	939	US-10-437-963-81637	Sequence 81637, A
28	184	7.7	274	US-10-389-048-16407	Sequence 16407, A
29	159	6.7	921	US-10-424-599-108137	Sequence 108137, A
30	153	6.4	3209	US-10-170-235-21113	Sequence 21113, A
31	140.5	5.9	350	US-10-424-599-9668	Sequence 9668, App
32	138	5.8	477	US-09-837-604B-52491	Sequence 52491, A
33	138	5.8	477	US-09-837-604A-52491	Sequence 52491, A
34	132	5.6	2250	US-10-369-493-45620	Sequence 45620, A
35	128	5.4	405	US-10-437-963-13671	Sequence 13671, A
36	124	5.2	3931	US-10-437-963-13671	Sequence 13671, A
37	123.5	5.2	564	US-10-437-963-7175	Sequence 7175, App
38	122	5.1	441	US-10-437-963-98795	Sequence 98795, A
39	119	5.0	3282	US-10-149-310-215	Sequence 215, App
40	119	5.0	3282	US-10-149-310-215	Sequence 215, App
41	118	5.0	782	US-10-304-095-5	Sequence 3351, A
42	115.5	4.9	10640	US-10-304-095-5	Sequence 5, Appl1
43	115	4.8	2718	US-10-369-493-45577	Sequence 45577, A
44	114.5	4.8	2637	US-10-282-122A-16387	Sequence 16387, A
45	114	4.8	2892	US-10-282-122A-16387	Sequence 255, App

ALIGNMENTS

RESULT 1
US-10-425-114-11348
Sequence 11348, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425, 114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 11348
LENGTH: 1722
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 701055296_FLI
US-10-425-114-11348

Alignment Scores:
Pred. No.: 1.37e-116
Score: 1166.00
Percent Similarity: 65.65%
Best Local Similarity: 52.95%

Length: 1722
Matches: 242
Conservative: 58
Mismatch: 111

Query Match: 49.03% Indels: 46
 DB: 10 Gaps: 9
 US-09-890-220-2 (1-445) x US-10-425-114-11348 (1-1722)

```

QY 1 MetcysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 164 ATGTGCGGCGAAATTCCTCCGATACACATCCGGTGAAGAAAGAAATTCAGACTATAG 223
QY 21 AsnLeuLeuIleTyrcysLysProValArgLeuTyraAsnIlePheHisLeuArgSerLeu 40
DB 224 AGCCTTTGATTATATGCAAGCCGTGTGAACCTTTAAATAATTCCTGACCGCGCTCTTT 283
QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrlsIleGlyAlaLysArgLysArg 60
DB 284 CAAATCCTCTTTCTTAGAGATGTTCCGTTATTAATAATGAAGCAAGCCGTAAAGG 343
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrlsAspCysAsnAsnThrLeuGln 80
DB 344 AGCTTGAGAGCA--GGAAATGTGATTTCAATTAATAGGACCGCTACACATCTCTTGA 400
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 401 AAGACTGAGAGCAAGCAAGCACTTTCTTCGCGTTTGGCTGATGACAGTGGAAATCTT 460
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 461 AAGGTTTGGGATTTCTCTTTGTCATCATCATCATCATCATCATCATCATCATCATCAT 520
QY 121 PheGluGluTyrglnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 521 ACTGAGATTAACCAAGAGTGAATGATCTCTGTAATAATGAT-----ATATGAGATCA 574
QY 141 GluIleSerAspSer-----AspLysPheGluProPheSerLeuGlySerLysPro 157
DB 575 GAGAAATGCTGCTGAGAGTGAATTCACAAATCGCAACCTTTCTTCTGCTGTAAGACT 634
QY 158 ArgLysArgArgGlnArgGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
DB 635 CGAAACAGTGAAGAAAGAAAGCACTGTTCAATTAATAAGCAACCAATTAATAATCTCG 694
QY 178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAsp 197
DB 695 GAGTTGAGCTACCA-----GAAGCATACATTAATGAATTTCTACAAAGATGAT 745
QY 197 ----- 197
DB 746 GATATCTTATCTGCAAGAGAGAAATGTGTCTAGAACATCTCTAGTGAAGATTTTC 805
QY 198 -----GlyAsnArgGlyLeuGlyTyrlsProGluAlaThrGluLeu 210
DB 806 CCTAGTGAAGAAATGACGAGGAGAAATTTGGTCTGATCATCTGACCATGAGCAAC 865
QY 211 AlaGlyIlePheGluMetThrSerAsnIleProPro-----AlaIleAlaHisSer 227
DB 866 CTGAGCATGTGGAATTCACATTTCCAGGTTTCAATTTGCCATGCCCAATCT 925
QY 228 SerLeuAspAlaGly-----AlaLysValIleLeuThrSerGluAlaValProAlaThr 246
DB 926 TCTGTGAGCCCTGAATGTGTAATCAATATGTAAAGTGAATCGCTGCTGCTGCT--- 982
QY 247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuLys 266
DB 983 AAAACAAAGAAAGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042
QY 267 ArgGlnPheTyrlsHisSerHisArgValGlnProMetAlaLeuGluGluValMetSerAsp 286
DB 1043 AGACTTTCTTTTCACTACACAGATGCAAGCTATGCGACTAGAAACAAGTGTTCACAC 1102
QY 287 ArgAspSerGluAspGluValAspAspValAlaAspPheGluAspArgGluMetLeu 306
DB 1103 CGTCAATGAGAAAGCAAGTGTGATGACGATTCGAGATCTTGAAGATAGAAAGATGCTT 1162

```

```

QY 307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPhe 326
DB 1163 GACGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1222
QY 327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSer 346
DB 1223 ATGAGAAAGCAAGGCTCTGACAGATGATGATGATGATGATGATGATGATGATGATGAT 1282
QY 347 ArgPheTyrgLysGluLeuHisArgTyrlsSerSerLeuPheProCysTyrlsArgLeuPhe 366
DB 1283 AAGCTTCATGAAAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1342
QY 367 LeuIleLysLeuThrPheAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
DB 1343 ATGATCAAAACCTTTGGAATCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1402
QY 387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerVal 406
DB 1403 GTATTAGATAGTATACAGAAATGAGGATCGGCTACAGAAATAT----- 1447
QY 407 AspArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisPro 423
DB 1448 TGAAGCAATTAAGAAAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1498

RESULT 2
US-10-424-599-111422
; Sequence 111422, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: Ia Rosa Thomas J
; APPLICANT: Kowalik David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111422
; LENGTH: 1838
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1
US-10-424-599-111422

Alignment Scores:
Pred. No.: 1,52e-116 Length: 1838
Score: 1166.00 Matches: 242
Percent Similarity: 65.65% Conservative: 58
Best Local Similarity: 52.95% Mismatches: 111
Query Match: 49.03% Indels: 46
DB: 10 Gaps: 9
US-09-890-220-2 (1-445) x US-10-424-599-111422 (1-1838)
QY 1 MetcysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 164 ATGTGCGGCGAAATTCCTCCGATACACATCCGGTGAAGAAAGAAATTCAGACTATAG 317
QY 21 AsnLeuLeuIleTyrcysLysProValArgLeuTyraAsnIlePheHisLeuArgSerLeu 40
DB 224 AGCCTTTGATTATATGCAAGCCGTGTGAACCTTTAAATAATTCCTGACCGCGCTCTTT 377
QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrlsIleGlyAlaLysArgLysArg 60
DB 378 CAAATCCTCTTTCTTAGAGATGTTCCGTTATTAATAATGAAGCAAGCCGTAAAGG 437
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrlsAspCysAsnAsnThrLeuGln 80
DB 438 AGCTTGAGAGCA--GGAAATGTGATTTCAATTAATAGGACCGCTACACATCTCTTGA 494
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100

```

[illegible]

Db 1542 TGAGACCATTAAGAAAGAACTGACTCAACAAAGAAATATACAGCTTCTCCG 1592

RESULT 3
US-10-425-114-31883
: Sequence 31883, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jindong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E.
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53313)B
: CURRENT APPLICATION NUMBER: US/10/425,114
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 31883
: LENGTH: 1841
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-ZMFLB73226G12.FL1
US-10-425-114-31883

Alignment Scores:
Pred. NO.: 4.12e-80 Length: 1841
Score: 833.00 Matches: 180
Percent Similarity: 60.81% Conservative: 45
Best Local Similarity: 48.65% Mismatches: 83
Query Match: 35.03% Indels: 62
DB: 10 Gaps: 8

US-09-890-220-2 (1-445) x US-10-425-114-31883 (1-1841)

QY 43 ProSerPheLeuProArgCysLeuAsnTyrlYsIleGlyAlaIalysArgYlsArgYlsSer 62
Db 459 CCTTCTCTATTATCCGAAGATCATAGTTGACA----- 491

QY 63 ArgSerThrngIyMetValValPheAsnTyrlYsAspCysAsnAsnthrLeuGlnLysThr 82
Db 492 -----GCTGGAATGCTACTTTTATTATTCAGACTACTACAAATATCAATGCAAAAGACT 545

QY 83 GluValArgGluAspCysSerCysProPheCysSerMetLeuIscGlySerPheLysGly 102
Db 546 GAAGTACTGGAAGATTTTCTGCGCAATTTGCTGTCGTGCATGCAATGTGAAGCTTCAGAGGC 605

QY 103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 122
Db 606 CTAGAGATCCACTTAATTCATCATCATGACCTGTTCCATCATGACGTTGTGATATCTGAA 665

QY 123 GluTyrglnThrValAsnValSerValIysLeuAsnSerPhe-----IlePheGluGln 140
Db 666 GAGTCCCAAGGCTGTGAATGTTAGTGTGAAGGCTGTGAGCCTGGAAAACCTAGATTGTGCA 725

QY 141 GluGlySerAspAspLysPheGluProPheSerLeuIscSerLysProArgLysArg 160
Db 726 GAGGAGAGT---GATCCAAAGCATCAACATTTTCTACTGCTCGAGGCTTAAAGAGCGT 782

QY 161 ArgGlnArgGlyGlyArgAsnAsnthrArgArgLeuLysValCysPheLeuProLeuAsp 180
Db 783 AGA-----AGGTTTGGAAACCAACAGCTGAGAAA-----TTCAAGCATGTTTAC 824

QY 181 SerProSerLeuThrAsnGlyThr----- 188

Db 825 CCACATATTATGGAATCAGAGATCACATGAAAGATGCCCAAGGAGGCTCTGAAGATGACTAT 884

QY 189 -----GluAsnGlyIleThrLeuAsnAsnAspGlyAsnArgGlyLeuGlyTyPro 205
Db 885 GTGCAAGAGCAAAATGGGCTTCTGTAGCAAAAT----- 917

QY 206 GluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProAlaIleAla 225

RESULT 5

US-10-437-963-91042
; Sequence 91042, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 91042
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_89656C.1
US-10-437-963-91042

Alignment Scores:

Pred. No.: 1,13e-59 Length: 2700
Score: 648.50 Matches: 168
Percent Similarity: 36.23% Conservative: 70
Best Local Similarity: 25.57% Mismatches: 104
Query Match: 27.27% Indels: 315
DB: Gaps: 11

US-09-890-220-2 (1-445) x US-10-437-963-91042 (1-2700)

OY 1 Metcysarglnasnancysarqalalysserproclugluvalilieserthrpsclu 20
||| ||| : : : : | : : : : | : : : : | : : : : |
Db 516 ATGCGTGTGGACCTGTGCCTCACTTCCTTCATCAACAACCTTAACCCCTGAAGAA 575
||| ||| : : : : | : : : : | : : : : | : : : : |
OY 21 Asnleuleuiletyrcylsyrprovalargleuyrnsnllephehtisleuarserlu 40
||||| : : : : | : : : : | : : : : | : : : : |
Db 576 AAATTAGCTTAATCTGCACGCCACTTGAGTTGACAACTTATTTCACACCAGCCATT 635
||||| : : : : | : : : : | : : : : | : : : : |
OY 41 GlyasnProserphelenProArGcysleuasntyrlslllegialatalysarGlysarg 60
||||| : : : : | : : : : | : : : : | : : : : |
Db 636 GAANAATCCTCTTATCTTCAAAGAATGCTCTTATATAAGATAGCTGCAAAAACAAAAA 695
61 lys----- 61

Pred. No.: 6,13e-57 Length: 1004
 Score: 617.50 Matches: 114
 Percent Similarity: 76.92% Conservative: 26
 Best Local Similarity: 62.64% Mismatches: 39
 Query Match: 25.97% Indels: 3
 DB: 10 Gaps: 1

US-09-890-220-2 (1-445) x US-10-437-963-11609 (1-1004)

```

OY 223 AlaIleIaHisSerSerLeuAspAlaGlyAlaVal-----IleLeuThrSer 239
      ::::::::::::::::::::
DB 844 TCTGAGACACAGCTTGTGATCTGTGTAATTCATTCACGGTAGCAATCTTTCAGCA 785
OY 240 GluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGlyLysSerGluAlaArg 259
      ::::::::::::::::::::
DB 784 CCAACAGTGTACATTTGGGAGACAGAAAGCTGTCTGTGAACAGCTATCCCA 725
OY 260 SerHisLeuLeuGlnLysArgGlnPheThrHisSerHisArgValGlnProMetAla 279
      ::::::::::::::::::::
DB 724 ATTCGACAGCTCTACAAAAGCCAGCTTTCATTCACAGGGCTCAACCAATGGCA 665
OY 280 LeuGlnGlnValMetSerAspArgAspSerGluAspGlnValAspAspValAlaAsp 299
      ::::::::::::::::::::
DB 664 TTGGAGCAAGTTTCTCAGATCGTATGATGAATGAATGAATGATGATGATGATGAT 605
OY 300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlnLysPhe 319
      ::::::::::::::::::::
DB 604 TTGAGAGATGAAGAAATGCTTGATGATTTGTTGATGTTACAAAGACAGCAATCTTAT 545
OY 320 MetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspLysHisLeuSer 339
      ::::::::::::::::::::
DB 544 ATGCAATGTGGAAATTCATTTGTCGGAACAAAGGTTAGCTAGCGATGGCATATTC 485
OY 340 TrpAlaCysGluAlaPheSerArgPheThrGlyLysGlnLeuHisArgLysSerLeu 359
      ::::::::::::::::::::
DB 484 TGGGATGCGAAGCAATCTCGCAGTTTCATGGACAAAGACTGTCAAAATCCACTCTA 425
OY 360 PheTrpCysTrpArgLeuPheLeuLysLeuTrpAsnHisGlyLeuValAspSerAla 379
      ::::::::::::::::::::
DB 424 CTAATGCTGTGAGAGCTTTTATATGTCAAACCTCTGACACCAAGCTACTGATCGCCA 365
OY 380 ThrIleAsnAsnCysAsnThrIleLeuGlnAsnCysArgAsnSerSerAspThrThr 399
      ::::::::::::::::::::
DB 364 GCCATGAATGCTCGACACACAAATCTTGAAGGCTACCTGAAGCAAGCTCGATCCAAAG 305
OY 400 ThrAsn 401
      ::
DB 304 AAAAAT 299

```

RESULT 7

US-10-424-599-141143/c

; Sequence 141143, Application US/10424599

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 141143

; LENGTH: 1070

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1

US-10-424-599-141143

Alignment Scores:

Pred. No.: 2.66e-52 Length: 1070

Score: 575.50 Matches: 123
 Percent Similarity: 63.14% Conservative: 38
 Best Local Similarity: 48.24% Mismatches: 63
 Query Match: 24.20% Indels: 31
 DB: 10 Gaps: 6

US-09-890-220-2 (1-445) x US-10-424-599-141143 (1-1070)

```

OY 145 AspArgLysPheGlnProPheSerLeuCysSerLysProArgLysArgGlnArgGly 164
      ::::::::::::::::::::
DB 1032 GATCCAAAGCTCCAAACATTTTCTTGTGGAAAGCCCTTAACCGTAGCAACAGCA 973
OY 165 GlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeu 184
      ::::::::::::::::::::
DB 972 GACCAATCT-----TTGAAAATGCA-----GTGGCTTACAGCTTCCTTCT 928
OY 185 ThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGly 204
      ::::::::::::::::::::
DB 927 GCAGAGGGAGCTGAT-----ATTGAGAGAGAGATGATGCT----- 892
OY 205 ProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProAlaIle 224
      ::::::::::::::::::::
DB 891 -----ATTTCGCCAAT 877
OY 225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThr-----SerGluAlaVal 242
      ::::::::::::::::::::
DB 876 ATTCGATCAGCTCCGATGAGACATCTGTTCAGTCAATCTCAGCTGATGATCAAGCA 817
OY 243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeu 262
      ::::::::::::::::::::
DB 816 CTTCAGTTTGGCAAGCAAGCAAGAGTTGCAATGAGCGCTGACCCAGCAACAGTACC 757
OY 263 LeuLeuGlnLysArgGlnPheThrHisSerHisArgValGlnProMetAlaLeuGln 282
      ::::::::::::::::::::
DB 756 TTCTTGAGAGACGCAAAATTTTCTCATTCACCAAAAGCTCAGCAATGCAATGCAACA 697
OY 283 ValMetSerAspArgAspSerGluAspGlnValAspAspValAlaAspPheGluAsp 302
      ::::::::::::::::::::
DB 696 GTTCATTCGCAATGAAGATGCGAAGATGAAGTGAATGATGATGATGATGATGATGAT 637
OY 303 ArgGlnMetLeuAspAspPheValAspValAsnLysAspGlnLysGlnPheMetHisLeu 322
      ::::::::::::::::::::
DB 636 CCAAGAGTCTTGAAGAAATGCTGTGATGACCAATGAGAGAGACTTTCATGATGAT 577
OY 323 TrpAsnSerPheValArgLysGlnArgValIleAlaAspGlnHisIleSerTrpLac 342
      ::::::::::::::::::::
DB 576 TGGAACTCATTTGTTGGAGAGCATGCTGTGATGACGATGATGATGATGATGATGAT 517
OY 343 GluAlaPheSerArgPheThrGlyLysGlnLeuHisArgLysSerSerLeuPheTrpCys 362
      ::::::::::::::::::::
DB 516 GAGGCTTCTCAAAATTCGATGACCTGAGTTTGTTCATCTCCCTCACTGCGACAGGCT 457
OY 363 TrpArgLeuPheLeuLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsn 382
      ::::::::::::::::::::
DB 456 TGGAGAAATATTATGTCATCAATTAACAATCAATGCTCTGATGATGCTGCGACATGAT 397
OY 383 AsnCysAsnThrIleLeuGlnAsnCys---ArgAsnSerSerAsp 396
      ::::::::::::::::::::
DB 396 GACTGTAAATATTATCTTGAACAATACCAAAAGCAAGAAATTCAGAT 352

```

RESULT 8

US-10-425-114-30709

; Sequence 30709, Application US/10425114

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53313)B

Fri Jun 20 09:03:41 2003

us-09-890-220-2.rtf

Page 8

```

CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 30709
LENGTH: 2439
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
US-10-425,114-30709
DC-ZMFLB73067D07_FIL

```

pred. NO.:	1,286-42	Length:	2439
Score:	497.00	Matches:	160
Percent Similarity:	29.74%	Conservative:	55
Best Local Similarity:	22.13%	Mismatches:	99
Query Match:	20.69%	Indels:	411
DB:	10	Gaps:	12

US-09-830-220-2 (1-445) x US-10-425-114-30709 (1-2439)

OY		3	ArgGlaAsnCyArqAlaIySerSerProGluGluValIleSerThrAspGluIleu	22
Dd		283	CGCAGACAGCTAAGAATCCGATTTGCTCCAGATGTGAACACTTACCCTGGTA AAAAGATTGG	342
OY		23	LemIeTyECyAlysProValaIrgLeuVTrasnllePheHisLeuAtySseLauGlyaaN	42
Dd		343	GCTGTGTATTGCAGACGACGATCGAGCTCAACAATTTATTCAACGCCACGACNTGAAAAAT	402
OY		43	ProserPheIlePrArCySleuSuSoLyrlsIleGlyAlaySaArqLySarY-----	60
Dd		403	CCCCCTTTATTCAAAGATGCCCTCTTTATATATACATGAGAGAGAAAAAGAGATT	462
OY		60	-----	60
Dd		463	CAGATAACCATATTCACCTTTCTGGAGTGACAAATACTAGTTGCAAAACACTTATGTCITT	522
OY		60	-----	60
Dd		523	CCCTTTATGTTCTGTAGACTAGAACCCACGTAACTTCCTCAC TTGAAGGCA TTTCTCCA	582
OY		60	-----	60
Dd		583	ATTATACGATTCAGTCGGGTYTGTGCTGTACTTCCTTTAGTAGAACATGSAATTAAGAC	642
OY		61	-----LysSerArqySerThr-----	65
Dd		643	AACAGTGAAGCTACACTTATCATCTTCTCTACGTTGAGAGATTGTCAACCTCCGCTGTGC	702
OY		65	-----	65
Dd		703	AACCATGATATTATCTTTATTAGCTGTGGGCAAGTTGGACAAGTAATGSGTAGATTAAC	762
OY		65	-----	65
Dd		763	TGCTCTGGAGAACCATGTGGAGAGATTCCTCTCCAAAATGCTTGAGGGAANAATCCTCGG	822
OY		65	-----	65
Dd		823	GGTAAANATACCACTRAATTTACTGTCTCACTCTTTGGAGAGTGTGTCAATTTAAGTTTG	882
OY		65	-----	65
Dd		883	GGACATATTGTGAGTGGCATCTPAAGTTACAAATGAGACCAAGCTTCTTAGAGCCAAA	942
OY		65	-----	65
Dd		943	TTTTGTGAGACAAGACGTTGCTTGACATTTTGTCTCAATAGGTTGATGCTGTGGGTCA	1002
OY		65	-----	65
Dd		1003	TATTAACATAACAATAATGATGCCGACAAAGAGCGCTGTGCAGAGATATGTCCTTGTGT	1062

Oy	6	-----		65
Db	1063	CCATATAGTAGTACTCATATATATGATGTCCACCTTCGTCATTATACAGATATCATTAAG	1122	
Oy	66	-----GlymetValValPheasnTyrltylaspyasasnphnleuglnlystr	82	
Db	1123	TTAAGATCTGGCAATGTCTTTTATATACAGACTACATATATCATGCAAGACT	1182	
Oy	83	GlulValArgGlulAspCysSerCysProPheCysSerMetLeucCysGlySerPheIysGly	102	
Db	1183	GAAGTACTGAAAGATTTCTCTGTCCATTGTGTGATATACAGATGTGAAGCTTCAGAGGT	1242	
Oy	103	LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGlu-PheIysLeuPhe-	121	
Db	1243	CTAGGATGGCCATTTTAAACCTCATCATAGATCTATTCCATATAGTTTTGGGTGATATC	1302	
Oy	122	-GluGlnTyrglnThrValAsnValSerValLysLeuAsnSerPhe-	136	
Db	1303	TGAAGAGTACCAAGGTGTTTAAATGTTAGTGTGAAGGCTGATCTTGGAGAAACAGAGTGTTC	1362	
Oy	137	-----IlePheGluGluGlnGly-	142	
Db	1363	TTCCAAAGGTTCGTTTAAATGAGAGCCTCCACTTGGTGGCAAAACCGAGGCATTTGATCAG	1422	
Oy	143	-----SerAspAspAsp-	156	
Db	1423	TTTCCTTAGTCACGAGACGACTCCGCTTGTGATTTCTTCCTTCTTGAG-TGCAG-NA	1480	
Oy	156	SP-DA-Y- 	158	
Db	1481	GCCAAAGATGAAGAAGCATTTGGCTATTCCTATCTCATCTCCATCAAAATTCCTTACATAGTGTGA	1540	
Oy	158	-----	158	
Db	1541	TTGTGACTCCATATANTTTTGGTGATATGTTTACATGCTTAAACAATGCATGTGA	1600	
Oy	158	-----	158	
Db	1601	CTATTTCTTTAAATGGAAGACACTCTTAACTGTATTTGTACTAATGATCTTAAGTGA	1660	
Oy	158	-----	158	
Db	1661	GCTTTTGGCGAGGCGTGTGATCCAAAGCATCAAACTTTCTTATCGCTCAAGGTTTAA	1720	
Oy	159	---LysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuValCysPheLe	177	
Db	1721	GAAGCTACACCATCAAAAGAACACAAATGGAGAAAATCAGGACATGACACTCATATATTAT	1780	
Oy	177	uproLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAs	197	
Db	1781	GGATATCAGGTCTACCTGTGAAGATGAGGACAGATCTGAGGACAACTTTGTG-----CA	1831	
Oy	197	pcgIAsnArgGlyLeuGlnGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetTh	217	
Db	1832	AGGGGAGGATGG-----	1844	
Oy	217	rSerAsnIleProPheAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal----	235	
Db	1845	-----ACTCTGTAGCAAAATGCTTCGATGTATCTCGTCAATCTTTTCAATCG	1891	
Oy	236	---IleuThrSerGluAlaValValProAlaThrIleThrValArgLysLeuSerAlaG	254	
Db	1892	CAGCAATCTTTTACCACCAACAGATCTACAGTTTGGGAACACAGAAAGCTATCT--GA	1948	
Oy	254	uArgSerGluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheIleThrHisSerHisAr	274	
Db	1949	GGAGTCTGGACCTCAAAATCGGCAACCTCTGGCAAAAACGACACTTCTTCATCTTCCACAG	2008	
Oy	274	gValGlnProMetAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAs	294	
Db	2009	G-----	2009	
Oy	294	pAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLys	314	

[illegible]


```

: CURRENT APPLICATION NUMBER: US/09/531,113
: CURRENT FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 48629
: SEQ ID NO 5697
: LENGTH: 554
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: jC-gmfl02220073g03a1
US-09-531-113-5697

Alignment Scores:
Pred. No.: 6,64e-28 Length: 554
Score: 348.50 Matches: 76
Percent Similarity: 66.42% Conservative: 13
Best Local Similarity: 56.72% Mismatches: 26
Query Match: 14,668 Indels: 20
DB: 5 Gaps: 2

US-09-890-220-2 (1-445) x US-09-531-113-5697 (1-554)
QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 193 ATGTGCGCGCAAAATTCCTCCGGTACACACATGCCGGTGAAGAAGAAATTCACCTGATGAG 252
QY 21 AsnLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 253 ACTCTTTAATTATTGCAAGCCTGTGTAACGTGTACAAATTTCTACCGCGCGCTCTT 312
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 313 CAAATCCCTCTTCTTCTTAGAGAGATGTTGCTTTATAAATAAGACCAAGCGCTAAAGG 372
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 373 AAGTTGAGAGCA---GGAATGTGATTTCAATATAGGCACTACACACTTCTTCG 429
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 430 AAACGTGAAGTGCAGACCAACTTTCTTCCGTTTGTGTTGATGACAGTGTGCACAC--- 486
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGlyLeu 120
DB 487 -----TTTAGCTTA 495

QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
DB 496 CTGGAGAT-TACCAAGCAGTGAATGTCTCCGTGAATAATTAAT 536

RESULT 14
US-09-531-113-5697
: Sequence 5697, Application US/09531113
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: Heck, Gregory R.
: APPLICANT: La Rosa, Thomas J.
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: FILE REFERENCE: 38-21(15761)B
: CURRENT APPLICATION NUMBER: US/09/531,113
: CURRENT FILING DATE: 2000-03-22
: NUMBER OF SEQ ID NOS: 48629
: SEQ ID NO 5697
: LENGTH: 554
: TYPE: DNA
: ORGANISM: Glycine max
: OTHER INFORMATION: Clone ID: jC-gmfl02220073g03a1
US-09-531-113-5697

Alignment Scores:
Pred. No.: 6,64e-28 Length: 554
Score: 348.50 Matches: 76
Percent Similarity: 66.42% Conservative: 13

```

```

Best Local Similarity: 56.72% Mismatches: 26
Query Match: 14,668 Indels: 20
DB: 6 Gaps: 2

US-09-890-220-2 (1-445) x US-09-531-113-5697 (1-554)
QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 193 ATGTGCGCGCAAAATTCCTCCGGTACACACATGCCGGTGAAGAAGAAATTCACCTGATGAG 252
QY 21 AsnLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 253 ACTCTTTAATTATTGCAAGCCTGTGTAACGTGTACAAATTTCTACCGCGCGCTCTT 312
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 313 CAAATCCCTCTTCTTCTTAGAGAGATGTTGCTTTATAAATAAGACCAAGCGCTAAAGG 372
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 373 AAGTTGAGAGCA---GGAATGTGATTTCAATATAGGCACTACACACTTCTTCG 429
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 430 AAACGTGAAGTGCAGACCAACTTTCTTCCGTTTGTGTTGATGACAGTGTGCACAC--- 486
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGlyLeu 120
DB 487 -----TTTAGCTTA 495

QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
DB 496 CTGGAGAT-TACCAAGCAGTGAATGTCTCCGTGAATAATTAAT 536

RESULT 15
US-09-837-604B-940/C
: Sequence 940, Application US/09837604B
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Shukla, Hridayabhiranjan
: APPLICANT: De La Pena, Robert C.
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: FILE REFERENCE: 38-21(51892)B
: CURRENT APPLICATION NUMBER: US/09/837,604B
: CURRENT FILING DATE: 2001-04-18
: PRIOR APPLICATION NUMBER: US 60/197,872
: PRIOR FILING DATE: 2000-04-19
: NUMBER OF SEQ ID NOS: 81288
: SEQ ID NO 940
: LENGTH: 531
: TYPE: DNA
: ORGANISM: Oryza sativa nipponbare
: FEATURE:
: OTHER INFORMATION: Clone ID: LIB3431-030-PI-NI-G11
US-09-837-604B-940

Alignment Scores:
Pred. No.: 9.44e-22 Length: 531
Score: 292.00 Matches: 49
Percent Similarity: 75.31% Conservative: 12
Best Local Similarity: 60.49% Mismatches: 20
Query Match: 12,288 Indels: 0
DB: 5 Gaps: 0

US-09-890-220-2 (1-445) x US-09-837-604B-940 (1-531)
QY 321 HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB 529 CATATGTGAATTCATTCTTGTGGAACAAAGGTTACTAGCGGATGCGCATATTCCTCG 470

```

Fri Jun 20 09:03:41 2003

us-09-890-220-2.rnpn

Page 12

```
QY 341 AlAcYsGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
Db 469 GCATGCGAAGCATTCCTCGACGATTTCATGCAAGAACTGTACAAAATCCAGCTCTACTA 410
QY 361 TTPCystTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
Db 409 TGGTGTGGAGGAGTTTATATGCAAACTCTGGAACCAAGTCTACTGATGCGGAGCC 350
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
Db 349 ATGAATGCTGCAACACAAATTCCTGAAGGCTACCTGACGGAAGCTCGGATCCAAAGAA 290
QY 401 Asn 401
Db 289 AAT 287
```

Search completed: June 20, 2003, 02:31:35
Job time : 982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:32:03 ; Search time 42 Seconds
(without alignments)
1018.567 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRKAKSSPEEVISTDE.....INNKNYDNKDNNSRDYIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	26.5	240	2 H71435	hypothetical prote
2	423	17.8	851	2 G71435	hypothetical prote
3	322.5	13.6	623	2 D71435	hypothetical prote
4	300	12.6	582	2 E84771	fertilization inde
5	266	11.2	295	2 A71436	hypothetical prote
6	147.5	6.2	4550	2 T18440	hypothetical prote
7	140.5	5.9	1711	2 T18449	hypothetical prote
8	133.5	5.6	3848	2 T17414	TIPC protein - sli
9	132	5.6	749	2 A57276	MAD1 protein - yea
10	131.5	5.5	1245	2 D71613	GAF domain protein
11	130.5	5.5	699	2 T18426	hypothetical prote
12	128.5	5.4	1817	2 D71606	hypothetical prote
13	128.5	5.4	3844	2 T18402	asparagine/asparta
14	127.5	5.4	448	2 S05355	hypothetical prote
15	127.5	5.4	947	2 T08605	hypothetical prote
16	127.5	5.4	961	2 S67368	probable membrane
17	127.5	5.4	962	2 JCS808	G protein-coupled
18	126.5	5.3	1584	2 T18276	protein-tyrosine k
19	126.5	5.3	2150	2 S71629	sensory transducti
20	126	5.3	518	2 S23692	erythrocyte membra
21	125.5	5.3	1402	2 T17456	cell surface prote
22	124.5	5.2	686	2 A71607	Mtn3/PAG1P-like p
23	124	5.2	1619	2 T18499	hypothetical prote
24	123.5	5.2	1858	2 T18273	1-phosphatidylinos
25	123.5	5.2	2010	2 B71616	phosphatase (acid
26	123.5	5.2	3394	2 T18501	hypothetical prote
27	122.5	5.2	2269	2 T18472	hypothetical prote
28	122	5.1	720	2 T51007	hypothetical prote
29	122	5.1	1457	2 T14577	protein kinase Yak

30	122	5.1	1714	2 E71609	Ser/Thr protein k1
31	120.5	5.1	608	2 T18437	hypothetical prote
32	120.5	5.1	1436	2 D71618	hypothetical prote
33	120	5.0	911	2 T18451	hypothetical prote
34	120	5.0	1119	2 T18491	hypothetical prote
35	120	5.0	1195	2 S38174	hypothetical purine nu
36	120	5.0	2380	2 E71604	hypothetical prote
37	119	5.0	1094	2 S46021	probable regulator
38	118.5	5.0	1278	2 A71609	probable secreted
39	118	5.0	3724	2 T18427	hypothetical prote
40	117.5	4.9	974	1 A40213	optic lobe develop
41	117.5	4.9	1075	2 S54067	probable membrane
42	117	4.9	590	2 S66956	hypothetical prote
43	117	4.9	2175	1 S03170	homeotic protein c
44	114.5	4.8	858	1 A42239	adenylate cyclase
45	114.5	4.8	1338	2 T18287	protein-tyrosine k

ALIGNMENTS

RESULT 1					
H71435	hypothetical protein - Arabidopsis thaliana				
C:Species: Arabidopsis thaliana (mouse-ear cress)					
A:Variety: Columbia					
C>Date: 03-Aug-1998	#sequence_revision 03-Aug-1998	#text_change 11-Jan-2002			
C:Accession: H71435					
R:Beyan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D					
.P.; Wedler, H.; Wedler, E.; Wandutt, R.; Wellzenegger, T.; Pohl, T.M.; Terry, N.; G					
avanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B					
Nature 391, 485-488, 1998					
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen					
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;					
C.; Chalatzis, N.					
A>Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t					
A:Reference number: A71400; MUID:98121113; PMID:9461215					
A:Accession: H71435					
A>Status: preliminary; nucleic acid sequence not shown; translation not shown					
A:Molecule type: DNA					
A:Residues: 1-240 <BEV>					
A:Cross-references: GB:Z97342; NID:92245031; PID:92245039					
C:Genetics:					
A:Map position: 4COP9-4G3845					
Query Match					
Best Local Similarity 47.6%; Pred. No. 1.1e-41;					
Matches 140; Conservative 2; Mismatches 0; Indels 152; Gaps 2;					
OY	4	ONCRKSSPEEYISTDENLTYCKVRLYNIFHLRSLGNPSFLPCLNKYGAKRRK--	61		
DB	51	ONCRKSSPEEYISTDENLTYCKVRLYNIFHLRSLGNPSFLPCLNKYGAKRRKYYA	110		
OY	62	-----SRSTGMVFNKYKDCNNLTOKTEVR	85		
DB	111	FLIECSCHSDMLFILLILMKLMLNYLSLSRYGHNSRSTGMVFNKYKDCNNLTOKTEVR	168		
OY	86	EDCSCPFCSMLCGSKGIQFHLNSHDFFERFLFEYQYVNVSVKLSNFIPEEGSD	145		
DB	169	-----	168		
OY	146	DKFEPFLSKSPRRKRGGRGNNTFRLLKVCFLPDPSPSLTNGTENGITLLDNGNGLGYP	205		
DB	169	-----GNRGIGTP	176		
OY	206	EATEIAGOFEMTSNIPPAIAHSSLDAGAKVLTSEAVVPATKRLKLSERSEAR	259		
DB	177	EATEIAGOFEMTSNIPPAIAHSSLDAGAKVLTSEAVVPATKRLKLSERSEAR	230		
RESULT 2					
G71435	hypothetical protein - Arabidopsis thaliana				

C:Species: Arabidopsis thaliana (mouse-ear cress)
 A:Variety: Columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
 C:Accession: G71435
 R:Byers, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Ditz
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Welzenegger, T.; Pohl, T.M.; Terry, N.; Giel
 Avangh, T.; Hempel, S.; Kottler, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 483-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
 erhoft, A.; Moore, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; Ans
 C.; Chaitwalis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
 A:Reference number: A71400; MID:98121113; PMID:9461215
 A:Accession: G71435
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-851 <BNV>
 A:Cross-references: GB:297342; NID:92245031; PID:6327517; PID:92245038
 C:Genetics:
 A:Map position: 4COP9-4G3845

Query Match 17.8%; Score 423; DB 2; Length 851;
 Best Local Similarity 39.6%; Pred. No. 6, 5e-25;
 Matches 101; Conservative 2; Mismatches 0; Indels 152; Gaps 2;

43 PSELPCLNYKIGAKRRK-----SRS 64
 |||||
 701 PSELPCLNYKIGAKRRKRYAFLECSCHSDMLFTLLIWKLMNLSLSRYGHNEWSRS 760
 |||||
 65 TGVVENVYKDCNNTLOKTEVEDSCPCFSCMLCSFKGLQFHLNSHDLFEFEKLEFEY 124
 |||||
 761 TGVVENVYKDCNNTLOKTE----- 779
 |||||
 125 QTVVSVKLSFIIEEGSDDKPEPFSLCSKPRRQGRGNTRRLKVCFLDPSPL 184
 |||||
 780 ----- 779
 |||||
 185 TNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHSILGAKYILNSEAVNP 244
 |||||
 780 -----GNGLGYPEATELAGOFEMTSNIPPAIHSILGAKYILNSEAVNP 826
 |||||
 245 ATKTRKLSAERSR 259
 |||||
 827 ATKTRKLSAERSR 841

RESULT 3
 D71435
 hypothetical protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 A:Variety: Columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
 C:Accession: D71435
 R:Byers, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Ditz
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Welzenegger, T.; Pohl, T.M.; Terry, N.; Giel
 Avangh, T.; Hempel, S.; Kottler, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 483-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
 erhoft, A.; Moore, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; Ans
 C.; Chaitwalis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
 A:Reference number: A71400; MID:98121113; PMID:9461215
 A:Accession: D71435
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-623 <BNV>
 A:Cross-references: GB:297342; NID:92245031; PID:92245035
 C:Genetics:
 A:Map position: 4COP9-4G3845

Query Match 13.6%; Score 322.5; DB 2; Length 623;
 Best Local Similarity 43.4%; Pred. No. 3, 1e-17;

Matches 76; Conservative 28; Mismatches 38; Indels 33; Gaps 6;

237 LTSEAVNPATRTKLSAERSRSHLT-LQRFYSHRVPALQVMSDRSDSEVD- 294
 |||||
 6 LTSEAVNPATRTKLSAERSRSHLT-LQRFYSHRVPALQVMSDRSDSEVD- 57
 |||||
 295 -DDVAFEDROMDDFDVYVK-DEKQPMHIMSFPKQRYANGHSMAEFSRFEKE 352
 |||||
 58 NDDAHLEESQMSNGMSNDENETVAERIKIMNFVQQRIVADAHIPMAEFSRLHGE 117
 |||||
 353 LHRYSLEWCWRLLFLIKLMNGLVDSATINNCITLDECRNSDPTTNNNSVD 407
 |||||
 118 LRSNLSL-----DLGNTEETETAPV-----ATEATTHQEDQDID 151
 |||||

RESULT 4
 E84771
 fertilization-independent seed 2 protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: E84771
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, E.; Tallon
 eus, D.; Mermann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MID:20083487; PMID:10617197
 A:Accession: E84771
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-582 <STO>
 A:Cross-references: GB:A8002093; NID:94263788; PIN:AAD15448.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g35670
 A:Map position: 2

Query Match 12.6%; Score 300; DB 2; Length 582;
 Best Local Similarity 19.6%; Pred. No. 1, 6e-15;
 Matches 117; Conservative 54; Mismatches 118; Indels 308; Gaps 13;

13 EVVISTDNLILYCKPRALNIFRLSGLPSFLPCLNYKIGAKRRKSRSTGVVNY 72
 |||||
 46 DDDVDVDENIITIKTPAVAVKLETRSKN----- 75
 |||||
 73 KDCNNTLOKTEVEDSCPCFSCMLCSFKGLQFHLNSHDLFEFEKLEFEYQ--TVVS 130
 |||||
 76 -----GLQHLKSSHDQEFKFEYRAEKDGPEDVS 106
 |||||
 131 VKLNSFIE--EGSDDKPEPFSLCSKPR-KRRQGRGNTRRLKVCFLDPSPLNG 187
 |||||
 107 VKSDTIKGVKRDVGNPQLSPFLFCCKNRNORRDBDNNVKKLVLLMELDLDLPRG 166
 |||||
 188 TENGITLLDNGNGLGYPEA-----TELAGEFEMTSNIP--PALAH 226
 |||||
 167 TENDSTHNDNVN--SPRAHSSEKISDILTQLAIAESSEKPVHNDGVNSPPRAH 225
 |||||
 227 SLSDGAVITLSEAVV----- 243
 |||||
 226 SBAEKNESTHVNDQVSSPPRAHSLKNESTHVNDNISPPKAHSSKKNSTHMD 285
 |||||
 244 ----PATRTKLSA-----ESESRLHLLQKQF-----YHSRVPMALE--- 281
 |||||
 266 VSPFRTSSKRETDILTQPAVSESEKVRVRRKQOLAKRYKAETQPAIAESSE 345
 |||||
 282 ----- 281
 |||||
 346 PKVILHVDENVSSPEAHSLEKASDILTQPAIAESSEPKVPHVDENVSTPRAHSSK 405
 |||||
 282 QVMSDRSDSEV-----DDVAD----- 299
 |||||
 406 KKSTRKKNVDVPSPPTRSKTSIDILTQPTIAESSEPKVRHVDNVSTPRAHSS 465
 |||||
 300 -----FEDRQ 304
 |||||

Db 466 KKKSTRKNDNIPSPKTRSSKTSNLTATQPAKAPSEPKVTRRLKGRQFYHSQWQ 525
 QY 305 MLDPEVDVNDKEQKFMHLMNSFVRKQRYIADGHISMACEAFREYKELHRSSLFW 361
 Db 526 RLRLVGVSKERKRYMILNIFVRKQRYIADGHVWACEEFAKLHKEKKNSSPDW 582

RESULT 5

A71436

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998

C:Accession: A71436; B71436

R:Bayan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk

F.; Medler, H.; Medler, E.; Wandul, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel

avanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech

erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans

C.; Chalwatis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal

A:Reference number: A71400; MUID:9812113; PMID:9461215

A:Accession: A71436

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-295 <BEV>

A:Cross-references: GB:297342; NID:92245031; PID:e327018; PID:92245041

A:Accession: B71436

A:Molecule type: DNA

A:Residues: 1-295 <BE2>

A:Cross-references: GB:297342; NID:92245031; PID:e327518; PID:92245040

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match

Best Local Similarity 11.2%; Score 266; DB 2; Length 295;

Matches 70; Conservative 3; Mismatches 30; Indels 28; Gaps 4;

QY 278 MLDPEVDVNDKEQKFMHLMNSFVRKQRYIADGHISMACEAFREYKELHRSSLFW 318

Db 1 MLDPEVDVNDKEQKFMHLMNSFVRKQRYIADGHISMACEAFREYKELHRSSLFW 60

QY 319 FMHLMNSFVRKQRYIADGHISMACEAFREYKELHRSSLFW 378

Db 61 FMHLMNSFVRKQRYIADGHISMACEAFREYKELHRSSLFW 111

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

QY 379 ATINNCNTILE 389

Db 112 ALANVHTRE 122

Matches 81; Conservative 64; Mismatches 150; Indels 115; Gaps 18;

QY 66 GMVFNVDKCN-----TLQKTEVREDCSCFSCMGSLQFOLNSSHDLFEFEK 119

Db 1820 GLTYTNLKNNN 1860

QY 120 L-----FEEOYVNVSV--KLNSFIFEESDDEKFEPSLCSKPRKRGGRNTYRLK 173

Db 1861 LNTLNKKYKNNNINCFYHYNSFLHSFYS-----EYIELIK-BERKIDKEKTKRK 1915

QY 174 VCFPLDPSLTNGTEGNTLLDNGRGIGYEPATELAGQFEMTSNIPALHSSLDAGA 233

Db 1916 YCHOEDRDDNN 1967

QY 234 KYILSEAVVPAITKTRKLSAERSSEARSHLLQKRFYHSRQVPALEQVMSDSEDEY 293

Db 1968 DQIINED-----NKKRKINNLSHETNM-----SDHNKP-----KKKKNSEYQI 2011

QY 294 DDDV-----ADFEEDQMLDDEVD-----VNKDEKQFMHLMNSFVRKQRYIADGH 337

Db 2012 NDDTKKKEHKKHEIDHKQKQDNCNGIEKKQVSKSNISKYVHFPERKNNI----- 2067

QY 338 ISWACEAFSREYKELHRSSLFW 397

Db 2068 -----KDEHK-----KEYNE-----RNDKMECDMLQNKRDNDHN 2099

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

Db 2100 NNNNNNNDDNNND 2148

QY 398 TTTNNNSVD-----RPSDSNTNNNNIYDHPNDIINNKNVKNKNSRD 441

[illegible]

RESULT 8

tipc protein slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 15-Oct-1999
 C:Accession: T17414
 R:Steger, J.T.; laub, M.T.; Loomis, W.F.
 Submitted to the EMBL Data Library, July 1998
 A:Description: Interaction of tip genes in early Dictyostelium discoideum development.
 A:Reference number: Z18774
 A:Accession: T17414
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 13848 <STP>
 A:Cross-references: EMBL:AA079445; NID:g3420744; PID:g3420745; PIDN:AA031916.1
 A:Experimental source: Strain AX4
 C:Genetics:
 A:Gene: tipC
 A:Introns: 72/3
 C:Function:
 A:Description: required for tip formation

Query Match	5.6%;	Score 133.5;	DB 2;	Length 3848;
Best Local Similarity	19.7%;	Pred. No. 0.18;		
Matches 75: Conservative	47;	Mismatches 103;	Indels 155;	

Matches 75; Conservative 47; Mismatches 103; Indels 155; Gaps 12;

76 NNTLQKTEVREDSCSPFCSMLCGSEFKGLQFHLNSSHDLFEFFFKLFEEYQTVNVSVKLS 135

DB 2040 NNSNNQLPKEDSQ-----LQKYLSISLEKEGF---ILNDHKISSPIKLLS 2083

150 F I E E G S D D D K F E P S L C S K P K R K R Q R G G R N N I R L K V C F L P L D S P S L I N G T E N G I T L L 195

COPIES STATION - 10012

— — — — —

Table 1. Demographic characteristics of study population

3158 | : : | : || : : | | : : | |
 3159 KRTFTMSTCEVOTIMADNYCOOZD ZWVCSTWZATFTTBMU
 3201

303 BOMT.DDEV DVNKDEKOE MHT.WNSEYBROPVATC HTSWACFAESPEVEKET HBYXCSY EWC 363

```

Db      2202 EOVNNDLIKPADDOK----- 2216
        | : | : | :

```

QY 363 WRLFLIKLWNHGLVDSATINNNTIILENCRRSSDTTTNNNNNSVDRPSDSNTNNNIVDH 422

```
Db      2217 -----NNNNNKKNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2257
```

QY 423 PNDINNKNNVDNKKDNSRDK 442

Db 2258 NNNNNNNNNKNNNDNNNNIK 2277

RESULT 9

A57276
MAD1 protein - Yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein G3191; protein YGL086w
C:Species: *Saccharomyces cerevisiae*
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 21-Jul-2000
C:Accession: A57276; S64093; S48246
R:Hardwick, K.G.; Murray, A.W.
J: Cell Biol. 131, 709-720, 1995
A:Title: Mad1p, a phosphoprotein component of the spindle assembly checkpoint in budding yeast
A:Reference number: A57276; MUID:96042315; PMID:7593191
A:Accession: A57276
A:Molecule type: DNA
A:Residues: 1749 <N>
A:Cross-references: GB:U14632; NID:9551091; PIDN:AA01620.1; PID:9551092
R:Rieger, M.; Mueller-Auer, S.; Bruckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64093
A:Molecule type: DNA
A:Residues: 1749 <N>
A:Cross-references: EMBL:672608; NID:91322610; PID:e243297; PID:91322611; MIPS:YGL086
A:Experimental source: strain 5286C
C:Genetics:
A:Gene: SGD:MAD1
A:Cross-references: SGD:S0003054; MIPS:YGL086w
C:Function:
A:Map position: 7L
C:Keywords: cell division control; nucleus; phosphoprotein

Query Match 5.6%; Score 132; DB 2; Length 749;
Best Local Similarity 21.3%; Pred. No. 0.029;
Matches 87; Conservative 57; Mismatches 143; Indels 122; Gaps 18;

QY 89 SCPEFCMGLGS-----FKGLPHLNSHDLFEPEFKLFEFYQVNYSVKLS 135
DB 40 SSFPLESPGSPDVSTNGCSNRQIQALQFKLTLQNEVEIKQLQKQNIIEKKYKA 98
QY 136 FTFE-EGSDDKF-----EPFSICSPRRRRGRGNRTKLYKCFPLDSPSLTNGTE 169
DB 99 TIDELKALNDIKYIESNDKLEQELSKERSANSMDKQ---CIEELRT----- 147
QY 190 NGITLLDNGCKGLGPEPEKTELAGEFE-MTSNIIPPAIASHSLDAGAKVITLSEAVYPATKT 248
DB 148 ---TLQNKD-----LEMETLRQDYDLSKLVNQCCHFKLEAES----- 183
QY 249 RKLSEARSEANSHLLQ-----KQEFYSHRVQPMALRQVNSDSEDEVDVADAFED 302
DB 184 -----SHSLMKYEKEIKROSVDIKDLQ---HQVM-----EKDDLSVYKA 221
QY 303 RQMLDPEVDYVKDEKQFQMLHNSFVRKQRYAD-----GHISWACEAFSRPEKE 352
DB 222 SKMISHPNYSTEE-----FNEILTENKMIQDOVQYTKLEELANMOQANEL-----KK 269
QY 353 LHRYSLEFWCMRLFLIKLMN-----HGL---VYSATINNCN-----TILENCRNSD 396
DB 270 LKQSDQTSFPMKLEKELQNLKLSQHLVLESQYENIQLEIMIDKSKLTKWEIYNDSDDDD 329
QY 397 TTTTNNNSVDRPSDNTNNNNIYVDHPNDINNNKNNVNDKNNRQKVK 445
DB 330 NNVNNDNNNNKNDNNNDNNNDNNDSNNNNNNNNRFRKNNIRNRPETIR 378

RESULT 10

D71613
GAP domain protein (cyclic nt signal transduct.) PFB0510w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: M71613
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

Query Match	5.5%;	Score 130.5;	DB 2;	Length 699;
Best Local Similarity	25.6%;	Pred. No. 0.035;		

	Query Match	5.4%;	Score 128.5;	DB 2;	Length 1817;
	Best Local Similarity	17.5%;	Pred. No. 0.17;		
	Matches	83;	Conservative	79;	Mismatches 160; Indels 151; Gaps 18;
OY	25 YCRFVRLIYNIVHILSLGNSFLPRLCLNKLIAK-----RRKRSTG--MVVPNY 72	: : : : : : : :			
Dd	382 YMRIKIKLEHIFELTEGCYLEL-LPNRKITSNNIYTPIITLKLIKNTITSNDVIYIHND 439	- - - - -			
OY	73 KDCNNLTQKTEVRDSCSPFCMSLGS-----FKGLQHLN-----SSHDLFEEFK 119	: : : : : : : :			
Dd	440 RTLRIRYIK-----FLSYRCSNNIYYHIIFENDLYITLHKKIFINNDMDLIKHYKS 488	- - - - -			
OY	120 -----LEEYQTVNV---SVKLNSFIIEEGSDDKKEPSLSCKPRKRORGGRNRT 170	- - - - -			
Dd	489 TADHIIYLITLNKNEMNIYNTILLNNY-----ISLVKKKKNNNEKROSLFN- 532	: : : : : : : :			
OY	171 RLAKCFLLPLDSPSLTNGTENGIITLLDNQRGLGYPEATELAGOPEMTSNIPALAHSSLD 230	- - - - -			

[illegible]

RESULT 13

T18402
 asparagine/aspartate rich protein - malaria parasite (Plasmodium falciparum) (fragments)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 09-Jun-2000
 C:Accession: T18402
 R:Barale, J.C.; Cannelle, D.; Attal-Bonnefoy, G.; Dehoux, P.; Bonnefoy, S.; Ridley, R.; Infect. Immun. 65, 3003-3010, 1997
 A:Title: Plasmodium falciparum AAP1, a giant protein containing repeated motifs rich in
 A:Reference number: Z18929; MUID: 97378065; PMID: 9234746
 A:Accession: T18402
 A:Status: preliminary; translated from GB/EMBL/DDAJ
 A:Molecule type: mRNA
 A:Residues: 1-3844 <BAR>
 A:Cross-references: EMBL:Y08926; NID:e1154302; PID:e311455; PIDN:CAA70130.1
 C:Genetics:
 A:Gene: aarp1

Query Match	Best Local	Matches	Score	DB 2	Length	3844
Similarity	19.1%	Pred. No. 0.45				
85	Conservative	49	Mismatches	141	Indels	169
						Gaps 14
QY	30	RYINPHLSTLSPFLPRCLNYKIGAKRRKRSSTGMVFN-----YD	74			
Db	992	OLENVPKRGFHHPSFEY-LVNKYGYKTHMPRSNVALLIKNNNSYLDAWPLSAFND	104.5			
QY	75	CNNLTOKTEVREDCSCPPCS-----MCGSFKIDLOHNLMSHDLDFEPEFKLEFEEQVYN	128			
Db	1050	FOFANEKCIKEENSYYIGCSNEDKEVYNLSPFRVQ-----NLLEYFK-----NGL	109.7			
QY	129	VSVKXLSNPFEEEGSDDDKFERFELSCSPRRRGRGNTRRLKYKCFELDSPSLTN--	186			
Db	1098	VSILLIVELINDE-----FQRIEKKKEELEKEEMKAKMDVYGEDD-SLNDK	114.7			
QY	187	-GEN-----GITLNDNGSLGYRETELACGFEMTSNIPRALHSSLDGAKVYLS	239			
Db	1148	GGTNNNTCDGVVINEANRSSG-----MSQN-----VYI	117.9			
QY	240	EAVVPATKTRKLSARSEARSHLLLOKQRFHSHKRVPMALDEQMSDRSEDEVDNDVAD	299			
Db	1180	ESGINNNDVLYGEERTNND-----IYMDNDNDNDNDNDNDNDNDNDNDNDNDNDND	121.9			
QY	300	FEDRQMLDFDYVYNKDEQFHLNNSVYKQVYADGCHISMACEAFSFEYKEKLHRYSL	359			
Db	1220	DVDDDDND	123.6			
QY	360	FWCWRLFLIKLHNGLDVSAIINCNLTILEKCRNSSDPTTTNNNSVDRPSDSTNNNNI	419			
Db	1237	-----ND	127.4			
QY	420	YDHPNDINN--KNVYDKKNSRD	441			
Db	1275	DDND	129.8			

RESULT 14
505355
Hypothetical protein (clone AAC1) - slime mold (Dictyostellium discoideum) (fragment)
C/Species: Dictyostellium discoideum
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 29-Oct-1999
C/Accession: S05355
R/Shaw, D.R.; Richter, H.; Giorda, R.; Omachi, T.; Ennis, H.L.
Mol. Gen. Genet. 218, 453-459, 1989
A/Title: Nucleotide sequences of Dictyostellium discoideum developmentally regulated CDNN-related sequences
A/Reference number: S05355; MUID:90066348; PMID:2511421
A/Accession: S05355
A/Molecule type: mRNA
A/Residues: 1448 <SHA>
A/Cross-references: EMBL:X16522; NID:97173; PIDN:CAA34529.1; PID:97174

Cross-references: EMBL:X16522; NID:g7173; PIDN:CAA34529.1; PID:g7174

Query Match 5.4%; Score 127.5; DB 2; Length 448;
Best Local Similarity 35.7%; P-rod No. 0.034;
Matches 25; Conservative 20; Mismatches 24; Indels 1; Gaps 1;

QY	37	NHGLVDASATINNCOTLEECRNSPTTNNNNSVDRPSDSNTNNNNIYDHPIDINKNK	431
Dd	273	NNNNNNNSNNNNNNNNNNSSNNNNNN--NNSTNNNTNNNNNNNTNNMTNNNNNN	331
QY	432	VDDKDNNSRD	441
Dd	332	INNINNNTNN	341

```

RESULT 15
T08605
Hypothetical protein HELE - slime mold (Dictyostelium discoideum) (fragment)
C:Species: Dictyostelium discoideum
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08605
R:Locms, W.F.; Iranfar, N.
Submitted to the EMBL Data Library, August 1997
A:Reference number: Z16451
A:Accession: T08605
A:Status: preliminary; translated from GR/EMBL/DBDU
A:Molecule type: DNA
A:Residues: 1-947 <DO>
A:Cross-references: EMBL:AF019961, NID:g2425112, PTD:g2425113
A:Experimental source: strain AX4
C:Genetics:
A:Gene: hele

```

[illegible]

Search completed: June 13, 2003, 15:38:03
Job time : 46 secs